Università degli Studi di Padova

FACOLTÀ DI SCIENZE MM. FF. NN.
Corso di Laurea Specialistica in Matematica

Tesi di Laurea

Large Deviations for small noise Itô processes through a weak convergence approach.

Candidato:
Alberto Chiarini
Matricola 1015439

Relatore:
Dr. Markus Fischer

Anno Accademico 2011 – 2012

Università degli Studi di Padova

FACOLTÀ DI SCIENZE MM. FF. NN.
Corso di Laurea Specialistica in Matematica

Tesi di Laurea

Large Deviations for small noise Itô processes through a weak convergence approach.

Candidato:
Alberto Chiarini
Matricola 1015439

Relatore:
Dr. Markus Fischer

Anno Accademico 2011 – 2012

Contents

Notations

Here there follow the most used notations throughout the thesis.

$\mathbb{R}^n$	n-dimensional real Euclidean space;
$\mathbb{R}^{n\times m}$	the set of all $(n \times m)$ real matrixes;
$\mathbb{R}^+$	the set of positive real numbers;
$a \wedge b$	the minimum between the real numbers a and b;
$a \vee b$	the maximum between the real numbers a and b;
$C_x[0,T]$	the space of continuous functions $\phi : [0,T] \to \mathbb{R}^d$ such that $\phi(0) = x$;
$AC_x([0,T];R^d)$	the space of absolutely continuous functions $\phi : [0,T] \to \mathbb{R}^d$ such that $\phi(0) = x$;
$L^p([0,T];\mathbb{R}^d)$	the set of all Lebesgue measurable functions $\phi : [0,T] \to \mathbb{R}^d$ such that $\int_0^T \lvert\phi(s)\rvert^p \, ds < +\infty$;
$L^\infty([0,T];\mathbb{R}^d)$	the set of essentially bounded measurable functions $\phi : [0,T] \to \mathbb{R}^d$;
S_N	the set of functions $\phi \in L^2([0,T];\mathbb{R}^d)$ such that $\int_0^T \lvert\phi(s)\rvert^2 \, ds \leq N$;
$\mathcal{W}^d$	the space of all continuous functions $\phi \in \mathcal{C}([0,T];\mathbb{R}^d)$ where T is a fixed positive real number;
θ	the Wiener measure over the space $\mathcal{W}^d$;
$\mathcal{B}$	the Borel σ-algebra of $\mathcal{W}^d$
W_t	the standard m-dimensional brownian motion;
$\{\mathcal{B}_t^W\}$	the filtration $\sigma(W_t : 0 \leq s \leq t)$;
$\{\mathcal{G}_t\}_{t\geq 0}$	the augmentation of the filtration generated by the brownian motion W_t;
$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$	a filtered probability space;
$\mathbb{E}_{\mathbb{P}}[X]$	the expectation of the random variable X with respect to $\mathbb{P}$;

$\mathcal{M}^p[0,T]$	the set of all progressively measurable processes v such that $\mathbb{E}[\int_0^T \|v_s\|^p\, ds] < \infty$;
$\mathcal{M}_b^p[0,T]$	the set of processes $v \in \mathcal{M}^p[0,T]$ such that $\int_0^T \|v_s\|^p\, ds < C$ for almost all ω;
$\Lambda^p[0,T]$	the set of all progressively measurable processes v such that $\mathbb{P}(\int_0^T \|v_s\|^p\, ds < \infty) = 1$;

Introduction

The theory of large deviations is concerned with the study of probabilities of rare events. It is not difficult to understand why rare events are important, to be persuaded it suffices to think about a lottery where events like hitting a jackpot can have an enormous impact.

Principles of large deviations may be effectively applied in information theory and risk management; in physics, an application arises in Thermodynamics and Statistical Mechanics.

At this point a non expert reader may wonder what is meant by rare event and large deviations in mathematical terms. We postpone the formal definition to the first chapter, and we begin by considering a very simple example on a familiar territory to provide some motivation as to what this thesis is about.

Let $X_1, X_2, \dots, X_n$ be a sequence of i.i.d. random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathbb{R}$, mean $\mu \in \mathbb{R}$ and variance $\sigma^2 \in (0, +\infty)$. As usual we denote by $\mathbb{E}$ the expectation under the measure $\mathbb{P}$. Let us define the random variables $S_n = X_1 + \cdots + X_n$. Thanks to the law of large numbers it follows that $\mathbb{P}$-almost surely

$$\frac{1}{n} S_n \to \mu.$$

The central limit theorem states that

$$\frac{1}{\sigma\sqrt{n}}(S_n - \mu n) \to Z \sim N(0,1)$$

in law. Accordingly we have an asymptotic estimate of the probability that S_n differs from μn by an amount of order $\sqrt{n}$, specifically $\mathbb{P}(S_n - \mu n \geq a\sqrt{n}) \to \mathbb{P}(Z \geq a/\sigma)$. These deviations are "normal".

Here we are interested in the case that S_n differs from μn by an amount of order n; these deviations are "large". An example is provided by the events $\{S_n \geq (\mu + a)n\}$, $a > 0$, whose probability is easily seen to go to zero as $n \to \infty$. The question is: how fast? The theory states that the decay is exponential in n, hence

$$\limsup_{n \to +\infty} \frac{1}{n} \log \mathbb{P}\left[S_n \geq (\mu + a)n\right] \leq -I(a), \qquad a > 0.$$

The function $a \to I(a)$ is known as the rate function; the knowledge of the function I is crucial when we need a correct evaluation of integrals of exponential functionals of S_n as $n \to \infty$, in which rare events play a considerable role.

Usually, the theory of large deviations is formulated in terms of a limiting behavior of normalized logarithms of probabilities of certain events,

$$\frac{1}{n} \log \mathbb{P}(X^n \in A).$$

In the weak convergence approach, these probabilities are replaced by normalized logarithms of expectations of exponentials of continuous functions,

$$\frac{1}{n} \log \mathbb{E}\left[e^{-nh(X^n)}\right].$$

We will refer to the study of the asymptotics of normalized logarithms of such expectations as the Laplace principle, which in the case of Polish spaces is equivalent to the large deviation principle.

To throw a glimpse on the weak convergence approach, let us consider a family of random variables $\{X^\epsilon\}_{\epsilon>0}$ with values in a Polish space $\mathcal{X}$, for example, we may consider the solutions of a stochastic differential equation with small noise

$$dX^\epsilon_t = b_\epsilon(t, X^\epsilon_t)\, dt + \sqrt{\epsilon}\sigma_\epsilon(t, X^\epsilon_t)\, dW_t \qquad X^\epsilon_0 = x.$$

The sequence is said to satisfy the Laplace principle with rate function I if for all bounded and continuous functions h mapping $\mathcal{X}$ into $\mathbb{R}$

$$\lim_{\epsilon \to 0+} -\epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] = \inf_{x \in \mathcal{X}} \{h(x) + I(x)\}.$$

A crucial step in the weak convergence approach is to rewrite the left hand side of the above display as a variational formula over the set of probability measures on $\mathcal{X}$, which we denote by $\mathcal{P}(\mathcal{X})$. Specifically, if μ^ϵ is the distribution of X^ϵ over $\mathcal{X}$, then it is not difficult to prove that

$$-\epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] = \inf_{\gamma \in \mathcal{P}(\mathcal{X})} \left\{ \int_{\mathcal{X}} h\, d\gamma + \epsilon R(\gamma \| \mu^\epsilon) \right\} \qquad (0.1)$$

where $R(\cdot\|\cdot)$ denotes the relative entropy.

Next one would like to send ϵ to zero; the limit obtained will yield the rate function I. To this end, usually, it is convenient to derive a different formula starting from the one given on the right hand side of (0.1). In the example of small noise diffusions it can be proved, under certain hypotheses, that

$$-\epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] = \inf_v \mathbb{E}\left[\frac{1}{2}\int_0^T \|v_s\|^2\, ds + f(X^{\epsilon,v})\right] \tag{0.2}$$

where $X^{\epsilon,v}$ is the solution of the controlled stochastic equation

$$dX^{\epsilon,v} = b_\epsilon(t, X_t^{\epsilon,v})\, dt + \sigma_\epsilon(t, X_t^{\epsilon,v})_\epsilon v_t\, dt + \sqrt{\epsilon}\sigma_\epsilon(t, X_t^{\epsilon,v})\, dW_t.$$

Now we can carry out the passage to the limit exploiting the convergence in distribution of X^{ϵ_n,v_n} for suitable controls v_n.

The use of weak convergence methods may be convenient in many situations, for example one can avoid awkward discretization methods that are employed in other approaches. The method provides a routine which can be applied in a wide range of problems: we refer to Dupuis and Ellis [1997] for examples of applications of this approach.

A summary of this thesis is as follows. Chapter 1 deals with the basic definition of large deviation principle and with its connection with the Laplace principle in the special setting of Polish spaces and good rate functions. A short section about relative entropy is inserted since it represents a cornerstone in the weak convergence approach. The chapter is mainly based on Dupuis and Ellis [1997, chapter 1], and on Dembo and Zeitouni [1998, chapter 4.3].

In chapter 2 we state a representation formula for bounded functionals of the Brownian motion W on bounded times, and we derive a representation formula for a process X which is the strong solution to a stochastic differential equation driven by W. The representation formula for the Brownian motion is due to Bou and Dupuis [1997], though we cover all the details for the proof of representation formula (0.2).

In chapter 3 we shall see an application of the weak convergence approach and of the formula (0.2), given in chapter 2; specifically, we derive a large deviation principle for small noise Itô processes under minimal assumptions; for the proof we have followed the approach found in Bou and Dupuis [1997], though the obtained result is more general since we assume that the coefficients may depend on ϵ and on the whole trajectory. We show in details a couple of examples for which the assumptions are satisfied. Particularly important is the case of Itô processes with locally lipschitz coefficients and sublinear growth at infinity, which requires a localization argument in order to apply the general theorem; the result can be used to prove, as a corollary, a large deviations principle for small random perturbations of systems with memory [see Mohammed and Zhang, 2006], and to recover Schilder's theorem and Freidlin-Wentzell estimates as special cases. Finally, we state and prove a large deviation principle for positive diffusions when the diffusion term is Hölder continuous, for which a particular case is the CEV model [see Baldi and Caramellino, 2011].

Some preliminary material required to be able to follow the proofs in the thesis is provided in the Appendix.

1

Large Deviations Principle and Laplace Principle

1.1 Formal definition

In this work the Large deviation principle will be established in an equivalent form, that of the Laplace principle. Our main references for this chapter are Dupuis and Ellis [1997] and Dembo and Zeitouni [1998].

We introduce the general notion of a large deviation principle, briefly LDP. We give such a principle over a family of random variables indexed by $\epsilon > 0$, for it will be convenient for our purposes in the sequel.

Let us set down some notation. Throughout this chapter, $\{X^\epsilon\}_{\epsilon>0}$ is a family of random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in a Polish space $\mathcal{X}$, that is, a complete separable metric space, whose metric will be denoted by d.

Definition 1.1. *A function I is called a (good)* ***rate function on*** $\mathcal{X}$*, or simply (good) rate function, if I maps $\mathcal{X}$ to $[0, +\infty]$ and if I has compact sublevel sets, which means that for each $M < +\infty$ the sublevel set*

$$\{x \in \mathcal{X} : I(x) \leq M\}$$

is compact. For a subset $A \subset \mathcal{X}$ we define $I(A) \doteq \inf_{x\in A} I(x)$.

Remark 1.1. *It is almost immediate to check that a function having compact sublevel sets is lower semicontinuous and that it attains its infimum on any nonempty closed set.*

For the sake of simplicity, we shall consider only the case of a good rate function. In the general definition the assumption of having compact sublevel sets is dropped for the weaker assumption of lower semicontinuity.

We are now ready to define the classical concept of a large deviation principle.

Definition 1.2 (LDP). *Let I be a rate function on $\mathcal{X}$. The family $\{X_\epsilon\}_{\epsilon>0}$ is said to satisfy a* ***large deviation principle on $\mathcal{X}$ with rate function*** *I if the following two conditions hold:*

(i) ***upper bound****: for each closed subset $F \subset \mathcal{X}$,*

$$\limsup_{\epsilon\to 0+} \epsilon \log \mathbb{P}(X^\epsilon \in F) \le -I(F)$$

(ii) ***lower bound****: for each open subset $A \subset \mathcal{X}$,*

$$\liminf_{\epsilon\to 0+} \epsilon \log \mathbb{P}(X^\epsilon \in A) \ge -I(A)$$

It is not difficult to prove that if a family satisfies a large deviation principle with rate function I, then the function is unique.

Example 1.1. A class of large deviation results involves the empirical means of i.i.d. random variables taking values in $\mathbb{R}^d$. To be precise, consider the empirical means $\hat{S}_n \doteq \frac{1}{n}\sum_{j=1}^{n} X_j$, $n \in \mathbb{N}$, for i.i.d. d-dimensional random vectors X_i distributed with common law $\mu \in \mathcal{P}(\mathbb{R}^d)$. The logarithmic moment generating function associated with the law μ is defined by

$$\Lambda(y) \doteq \log \mathbb{E}\left[e^{\langle y, X_1\rangle}\right], \qquad y \in \mathbb{R}^d,$$

where $\langle y, x\rangle$ is the usual inner product in $\mathbb{R}^d$. Let $\bar{x} = \mathbb{E}[X_1]$.

When $\bar{x}$ exists and is finite, and $\mathbb{E}[|X_1 - \bar{x}|^2] < \infty$, then $\hat{S}_n \to \bar{x}$ in probability as $n \to +\infty$ by the weak law of large numbers. Cramr's theorem characterizes the rate of this convergence by the following convex rate function

$$\Lambda^*(x) \doteq \sup_{y\in\mathbb{R}^d}\{\langle y, x\rangle - \Lambda(y)\}$$

namely the Fenchel-Legendre transform of $\Lambda(y)$. In other words,

(i) **upper bound**: for each closed subset $F \in \mathbb{R}^d$

$$\limsup_{n\to+\infty} \frac{1}{n} \log \mathbb{P}(\hat{S}_n \in F) \le -\Lambda^*(F)$$

(ii) lower bound: for each open subset $A \in \mathbb{R}^d$

$$\liminf_{n\to+\infty} \frac{1}{n} \log \mathbb{P}(\hat{S}_n \in A) \geq -\Lambda^*(A)$$

A proof of this theorem in more general settings can be found in Dembo and Zeitouni [1998, see section 2.2].

Example 1.2. Another class of large deviation results involves the sample path of stochastic processes which may satisfy a stochastic differential equation depending on a parameter ϵ. Specifically, let X^ϵ denote a family of processes which converges to some deterministic limit as $\epsilon \to 0$. A natural question that arises is what is the rate of this convergence. A classical result in this setting is **Schilder's theorem** which studies the asymptotics of the rescaled Brownian motion.

Let W denote a standard d-dimensional Brownian motion and for any $\epsilon > 0$ consider the process $X^\epsilon = \{X_t^\epsilon : t \in [0, T]\}$ which satisfies the stochastic differential equation

$$dX_t^\epsilon = \sqrt{\epsilon}\, dW_t$$

with initial condition $X_0 = 0$. Clearly $X_t^\epsilon = \sqrt{\epsilon}\, W_t$. As $\epsilon \to 0$ the law of X^ϵ converges weakly to the Dirac measure which is concentrated at the zero function, actually $\|X^\epsilon\|_\infty \to 0$ almost surely. We would like to understand how fast the probability of X^ϵ being away from the zero function goes to zero.

X^ϵ takes values in the space $C([0, T]; \mathbb{R}^d)$, which is a Polish space when equipped with the supremum norm. Let $AC_0([0, T]; \mathbb{R}^d)$ be the space of all absolutely continuous functions vanishing at zero. Schilder's theorem states that $\{X^\epsilon\}_{\epsilon>0}$ satisfies the large deviation principle on the Polish space $C([0, T]; \mathbb{R}^d)$ with good rate function

$$I(\phi) \doteq \begin{cases} \frac{1}{2}\int_0^T \|\dot{\phi}\|^2\, dt & \text{if } \phi \in AC_0([0, T]; \mathbb{R}^d), \\ \infty & \text{otherwise.} \end{cases}$$

A proof of this theorem can be found in Dupuis and Ellis [1997], or in [Dembo and Zeitouni, 1998, see theorem 5.2]. An alternative proof of Schilder's theorem is given as a corollary in Section 3.3.

1.2 Equivalent formulation of the LDP

The first step in deriving an equivalent formulation of the large deviations principle passes through the important Varadhan integral Lemma [Varadhan, 1966].

Theorem 1.1 (Varadhan). *Assume that the family $\{X^\epsilon\}_{\epsilon>0}$ satisfies the large deviation principle on a Polish space $\mathcal{X}$ with good rate function I and let $h : \mathcal{X} \to \mathbb{R}$ be any continuous function. Assume further the tail condition*

$$\lim_{M\to\infty} \liminf_{\epsilon\to 0} -\epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}} 1_{\{h(X^\epsilon)\le -M\}}\right] = +\infty \tag{1.1}$$

Then

$$\lim_{\epsilon\to 0+} -\epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] = \inf_{x\in\mathcal{X}}\{h(x)+I(x)\}. \tag{1.2}$$

This theorem is the natural extension of Laplace's method for studying the asymptotics of certain integrals on $\mathbb{R}$. Indeed, let $\mathcal{X} = \mathbb{R}$ and assume for the moment that the law of X^ϵ possesses a density with respect to the Lebesgue measure with a form such that $d\mathbb{P}(X^\epsilon \in dx)/dx \approx e^{-I(x)/\epsilon}$. Then

$$\int_{\mathbb{R}} e^{-h(x)/\epsilon}\, d\mathbb{P}(X^\epsilon \in dx) \approx \int_{\mathbb{R}} e^{-(h(x)+I(x))/\epsilon}\, dx;$$

if we assume that $I(\cdot)$ and $h(\cdot)$ are twice differentiable, with $h+I$ convex and possessing a unique global minimum at some $\bar{x}$, then Taylor's theorem yields

$$h(x)+I(x) = h(\bar{x}) + I(\bar{x}) + \frac{(x-\bar{x})^2}{2}(h(x)+I(x))''|_{x=\xi}$$

where $\xi \in [\bar{x}, x]$. Therefore,

$$\int_{\mathbb{R}} e^{-h(x)/\epsilon}\, d\mathbb{P}(X^\epsilon \in dx) \approx e^{-(h(\bar{x})+I(\bar{x}))/\epsilon} \int_{\mathbb{R}} e^{-B(x)\frac{(x-\bar{x})^2}{2\epsilon}}\, dx.$$

where $B(\cdot) \ge 0$. The content of Laplace's method is that on a logarithmic scale the rightmost integral may be ignored.

Theorem 1.1 is a direct consequence of the following two lemmas.

Lemma 1.1. *If $h : \mathcal{X} \to \mathbb{R}$ is upper semicontinuous and the large deviations lower bound holds with rate function I, then*

$$\limsup_{\epsilon\to 0} -\epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] \le \inf_{x\in\mathcal{X}}\{h(x)+I(x)\}. \tag{1.3}$$

Proof. Fix a point $x \in \mathcal{X}$ and $\delta > 0$. Since $h(\cdot)$ is upper semicontinuous, there exists an open neighborhood G of x such that $\sup_{y\in G} h(y) \le h(x) + \delta$. Hence

$$\begin{aligned}
\liminf_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] &\ge \liminf_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[1_G(X^\epsilon) e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] \\
&\ge -h(x) - \delta + \liminf_{\epsilon\to 0+} \epsilon \log \mathbb{P}(X^\epsilon \in G) \\
&\ge -h(x) - \delta - I(G) \\
&\ge -h(x) - I(x) - \delta.
\end{aligned}$$

Recall that we have chosen $x \in \mathcal{X}$ and $\delta > 0$ arbitrarily, therefore after taking the supremum over $x \in \mathcal{X}$ and sending $\delta \to 0+$, we obtain

$$\limsup_{\epsilon \to 0+} -\epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] \leq \inf_{x \in \mathcal{X}} \{h(x) + I(x)\}.$$

□

Lemma 1.2. *If $h : \mathcal{X} \to \mathbb{R}$ is a lower semicontinuous function for which the tail condition* (1.1) *holds, and the large deviations upper bound holds with good rate function I, then*

$$\liminf_{\epsilon \to 0} -\epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] \geq \inf_{x \in \mathcal{X}} \{h(x) + I(x)\}. \tag{1.4}$$

Proof. Consider first a function h bounded from below, and let $M > 0$ be such that $\inf_x h(x) \geq -M$. For such functions the tail condition (1.1) clearly holds true. Fix $\alpha < \infty$ and $\delta > 0$ and let $\Psi_I(\alpha) \doteq \{x : I(x) \leq \alpha\}$ denote the compact sublevel set of the good rate function I. Since I and h are lower semicontinuous, for every $x \in \Psi_I(\alpha)$ we find an open neighborhood A_x of x such that

$$\inf_{y \in \overline{A_x}} I(y) \geq I(x) - \delta, \qquad \inf_{y \in \overline{A_x}} h(y) \geq h(x) - \delta \tag{1.5}$$

Since

$$\Psi_I(\alpha) \subset \bigcup_{x \in \Psi_I(\alpha)} A_x$$

and $\Psi_I(\alpha)$ is compact, one can extract a finite subcover, for example A_{x_i}, $i = 1, \ldots, N$. Therefore

$$\begin{aligned}
\mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] &\leq \sum_{i=1}^{N} \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}} 1_{\{X^\epsilon \in A_{x_i}\}}\right] + e^{\frac{M}{\epsilon}} \mathbb{P}\left(X^\epsilon \in \left(\bigcup_{i=1}^{N} A_{x_i}\right)^c\right) \\
&\leq \sum_{i=1}^{N} e^{-\frac{(h(x_i)-\delta)}{\epsilon}} \mathbb{P}(X^\epsilon \in \overline{A_{x_i}}) + e^{\frac{M}{\epsilon}} \mathbb{P}\left(X^\epsilon \in \left(\bigcup_{i=1}^{N} A_{x_i}\right)^c\right)
\end{aligned}$$

where the last inequality follows by (1.5). Finally, we apply the large deviations upper bound to the closed sets $\overline{A_{x_i}}$, and to $(\bigcup_{i=1}^{N} A_{x_i})^c \subset \Psi_I(\alpha)^c$,

$$\begin{aligned}
&\liminf_{\epsilon \to 0+} -\epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] \\
&\qquad \geq \min\left\{\min_i \{h(x_i) - \delta + I(\overline{A_{x_i}})\}, M + I\left(\left(\bigcup_{i=1}^{N} A_{x_i}\right)^c\right)\right\} \\
&\qquad \geq \min\left\{\min_i \{h(x_i) + I(x_i) - 2\delta\}, M + \alpha\right\} \\
&\qquad \geq \min\left\{\inf_{x \in \mathcal{X}} \{h(x) + I(x)\}, M + \alpha\right\} - 2\delta.
\end{aligned}$$

Thus, for any function h bounded from below, the lemma follows by taking the limits $\delta \to 0$ and $\alpha \to \infty$.

To treat the general case, it suffices to consider

$$h_M(x) = h(x) \vee (-M) \geq h(x)$$

and use the preceding result to show that for every $M < +\infty$

$$\begin{aligned}\liminf_{\epsilon \to 0+} -\epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] \\ \geq \inf_{x \in \mathcal{X}} \{h(x) + I(x)\} \wedge \liminf_{\epsilon \to 0} -\epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}} 1_{\{h(X^\epsilon) \leq -M\}}\right]\end{aligned}$$

The tail condition (1.1) completes the proof of the lemma by taking the limit $M \to +\infty$. □

We have just proved that the large deviations principle implies the validity of the limit relation (1.2). Since the converse holds true, it is convenient to restate property (1.2) in terms of a definition.

Definition 1.3. *Let $\{X^\epsilon\}_{\epsilon>0}$ be a family of random variables with values on $\mathcal{X}$, and I be a good rate function on $\mathcal{X}$. The family $\{X^\epsilon\}$ is said to satisfy the* ***Laplace principle on $\mathcal{X}$ with rate function I*** *if for all bounded continuous functions $h : \mathcal{X} \to \mathbb{R}$*

$$\lim_{\epsilon \to 0+} -\epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] = \inf_{x \in \mathcal{X}} \{h(x) + I(x)\}$$

We say that $\{X^\epsilon\}$ satisfies the ***Laplace principle lower bound*** *if*

$$\liminf_{\epsilon \to 0+} -\epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] \geq \inf_{x \in \mathcal{X}} \{h(x) + I(x)\},$$

for all bounded continuous functions h. We say that $\{X^\epsilon\}$ satisfies the ***Laplace principle upper bound*** *if*

$$\liminf_{\epsilon \to 0+} -\epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] \leq \inf_{x \in \mathcal{X}} \{h(x) + I(x)\}$$

for all bounded continuous functions h.

Remark 1.2. *If we evaluate the Laplace limit for $h \equiv 0$ we can observe that the infimum of the rate function I on $\mathcal{X}$ equals zero. Since I is a function with compact sublevel sets, there exists a point $\bar{x} \in \mathcal{X}$ such that $I(\bar{x}) = 0$.*

With the last definition in mind, we can restate Varadhan's integral lemma by saying that the large deviation principle implies the Laplace principle.

In the next theorem we prove that the converse holds true. In the proof which follow the assumption of goodness of the rate function will be of great importance. In some ways the equivalence of the Laplace principle and of the large deviation principle can be seen as an analogue of Portemanteau theorem, where the equivalence between the weak convergence of probability measures and certain limits involving closed and open sets is proved.

Theorem 1.2. *The Laplace principle implies the large deviation principle with the same rate function. In other words, if I is a good rate function on a Polish space $\mathcal{X}$ and $\{X^\epsilon\}$ satisfies*

$$\lim_{\epsilon \to 0+} -\epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] = \inf_{x\in\mathcal{X}}\{h(x)+I(x)\}$$

for all bounded continuous functions h, then $\{\mathcal{X}^\epsilon\}$ satisfies the large deviation principle on $\mathcal{X}$ with rate function I.

Proof. By assumption, given a rate function I on $\mathcal{X}$, for all bounded and continuous functions h

$$\lim_{\epsilon \to 0+} -\epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] = \inf_{x\in\mathcal{X}}\{h(x)+I(x)\}$$

We want to prove that for each closed set $F \subset \mathcal{X}$ the family X^ϵ satisfies

$$\limsup_{\epsilon \to 0+} \epsilon \log \mathbb{P}(X^\epsilon \in F) \le -I(F)$$

and that for any choice of an open set $A \subset \mathcal{X}$ the family X^ϵ satisfies

$$\liminf_{\epsilon \to 0+} \epsilon \log \mathbb{P}(X^\epsilon \in A) \ge -I(A)$$

Large deviation upper bound. For any given closed set F, we consider the function

$$\phi(x) \doteq \begin{cases} 0, & \text{if } x \in F \\ +\infty, & \text{if } x \in F^c \end{cases}$$

clearly ϕ is lower semicontinuous. Next we would like to approximate ϕ through a sequence of bounded continuous function, so that the Laplace principle is applicable. With this aim, we define for any $j \in \mathbb{N}$ the function

$$h_j(x) \doteq j(d(x,F) \wedge 1)$$

where $d(x,F)$ denotes the distance from x to F. Clearly h_j is bounded by j and Lipschitz continuous of constant j, hence continuous. Moreover $h_j \uparrow \phi$ as $j \to \infty$. Therefore,

$$\epsilon \log \mathbb{P}(X^\epsilon \in F) = \epsilon \log \mathbb{E}\left[e^{-\frac{\phi(X^\epsilon)}{\epsilon}}\right] \le \epsilon \log \mathbb{E}\left[e^{-\frac{h_j(X^\epsilon)}{\epsilon}}\right]$$

And so, by sending $\epsilon \to 0+$, we get through the Laplace principle:

$$\limsup_{\epsilon \to 0+} \epsilon \log \mathbb{P}(X^\epsilon \in F) \leq - \inf_{x \in \mathcal{X}} \{h_j(x) + I(x)\}.$$

To complete the proof of the large deviation upper bound we need to show that the right hand side of the above inequality goes to $-I(F)$ as $j \to \infty$.

Since $h_j \leq \phi$,

$$\inf_{x \in \mathcal{X}} \{h_j(x) + I(x)\} \leq \inf_{x \in \mathcal{X}} \{\phi(x) + I(x)\} = \inf_{x \in F} I(x) = I(F),$$

hence

$$\limsup_{j \to \infty} \inf_{x \in \mathcal{X}} \{h_j(x) + I(x)\} \leq I(F)$$

Next we have to prove the converse inequality. It is clear that if $I(F) = 0$ then there is nothing to prove. Thus, we assume $I(F) > 0$.

By construction $h_j = 0$ on F, therefore

$$\begin{aligned} \inf_{x \in \mathcal{X}} \{h_j(x) + I(x)\} &= \min \left\{ \inf_{x \in F} \{h_j(x) + I(x)\}, \inf_{x \in F^c} \{h_j(x) + I(x)\} \right\} \\ &= \min \left\{ I(F), \inf_{x \in F^c} \{h_j(x) + I(x)\} \right\}. \end{aligned}$$

We are done once we prove that

$$\liminf_{j \to \infty} \inf_{x \in F^c} \{h_j(x) + I(x)\} \geq I(F). \tag{1.6}$$

We argue by contradiction. Assume that (1.6) is not true. If $I(F) = \infty$, then there exists a real number $M > 0$ such that

$$\liminf_{j \to \infty} \inf_{x \in F^c} \{h_j(x) + I(x)\} < M.$$

If $I(F) < \infty$, we just take $M \doteq I(F)$. Let us fix $0 < \epsilon < M/2$, then there exists a subsequence of $\{j\}_{j \in \mathbb{N}}$ such that for all indexes all j in this subsequence

$$\inf_{x \in F^c} \{h_j(x) + I(x)\} \leq M - 2\epsilon.$$

Moreover, for each j there exists $x_j \in F^c$ such that

$$h_j(x_j) + I(x_j) \leq M - \epsilon.$$

We claim that $d(x_j, F) \to 0$ as $j \to \infty$. Indeed, otherwise we can find a subsequence of $j \in \mathbb{N}$ and $\delta > 0$ such that $d(x_j, F) \geq \delta$ for any such j. Along that subsequence we have $h_j(x_j) \doteq j(d(x_j, F) \wedge 1) \geq j(\delta \wedge 1) \to +\infty$

as $j \to \infty$. Since $h_j(x_j) + I(x_j) \leq M - \epsilon < +\infty$, we have a contradiction. Being F closed, for any $j \in \mathbb{N}$ there exists $y_j \in F$ such that $d(x_j, F) = d(x_j, y_j)$; by the previous claim, $d(x_j, y_j) \to 0$ as $j \to \infty$. Observe that $\sup_{j \in \mathbb{N}} I(x_j) \leq M - \epsilon$, thus $\{x_j\}$ is contained in the compact sublevel set $\{x \in \mathcal{X} : I(x) \leq M - \epsilon\}$ (In fact, the rate function is good by assumption). Then, taking a further subsequence if necessary, we may assume that there exists $\bar{x} \in \{x \in \mathcal{X} : I(x) \leq M - \epsilon\}$ such that $d(x_j, \bar{x}) \to 0$ as $j \to \infty$. Through the triangular inequality we get $d(y_j, \bar{x}) \leq d(y_j, x_j) + d(x_j, \bar{x}) \to 0$. Now recall that F is closed, hence $\bar{x} \in F$, and $I(\bar{x}) \geq I(F) \geq M$. This last statement contradicts the hypothesis that $\bar{x} \in \{x \in \mathcal{X} : I(x) < M - \epsilon\}$. The large deviation upper bound is proved.

Large deviation lower bound. Let A be an open set. The statement is obviously true when $I(A) = \infty$, hence we can that $I(A) < \infty$. Let $x \in A$ be any point such that $I(x) < \infty$. Since A is open, we can find $\delta > 0$ such that the open ball $B(x, \delta)$ of center x and radius δ, is entirely contained in A. Define the function

$$h(y) \doteq M \left(\frac{d(y, x)}{\delta} \wedge 1 \right),$$

where M is a fixed real number greater than $I(x)$. The function h is obviously positive, bounded by M and continuous; in addition we have $h(x) = 0$, and $h(y) = M$ whenever $y \in B(x, \delta)^c$. Next we use the Laplace lower bound with the function h. We have

$$\begin{aligned} \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] &\leq e^{-\frac{M}{\epsilon}} \mathbb{P}(X^\epsilon \in B(x, \delta)^c) + \mathbb{P}(X^\epsilon \in B(x, \delta)) \\ &\leq e^{-\frac{M}{\epsilon}} + \mathbb{P}(X^\epsilon \in B(x, \delta)), \end{aligned}$$

Taking the lim inf on both sides, we get

$$\begin{aligned} \max\left\{ \liminf_{\epsilon \to 0} \epsilon \log \mathbb{P}(X^\epsilon \in B(x, \delta)), -M \right\} &\geq \liminf_{\epsilon \to 0+} \epsilon \log \mathbb{E}\left[e^{-\frac{h(X^\epsilon)}{\epsilon}}\right] \\ &\geq -\inf_{x \in \mathcal{X}} \{h(x) + I(x)\} \\ &\geq -h(x) - I(x) = -I(x). \end{aligned}$$

Recall that $M > I(x)$ and that $B(x, \delta) \subset A$; accordingly $-M < -I(x)$ and $\mathbb{P}(X^\epsilon \in A) > \mathbb{P}(X^\epsilon \in B(x, \delta))$. Therefore

$$\begin{aligned} \liminf_{\epsilon \to 0} \epsilon \log \mathbb{P}(X^\epsilon \in A) &\geq \liminf_{\epsilon \to 0} \epsilon \log \mathbb{P}(X^\epsilon \in B(x, \delta)) \\ &\geq -I(x) \geq -\inf\{I(x) : x \in A\} = -I(A). \end{aligned}$$

This ends the proof of the lower bound. □

Remark 1.3. *If we look at the preceding proof, we can observe how the Laplace principle upper bound implies the large deviation lower bound and the Laplace principle lower bound implies the large deviation upper bound.*

Remark 1.4. *In the proof of the upper bound we have seen in particular that, if we put $F = \{\xi\}$, then*

$$\lim_{j\to\infty} \inf_{x\in\mathcal{X}} \{h_j(x) + I(x)\} = I(\xi),$$

where $h_j(x) = j(d(x,\xi) \wedge 1)$ as above.

The above remark leads to an immediate proof of the uniqueness of the rate function I via the Laplace principle.

Theorem 1.3. *Assume that $\{X^\epsilon\}$ satisfies the Laplace principle on $\mathcal{X}$ with rate function I and with rate function J. Then $I(\xi) = J(\xi)$ for all $\xi \in \mathcal{X}$.*

Proof. For $j \in \mathbb{N}$ and $\xi \in \mathcal{X}$ consider $h_j(x) = j(d(x,\xi) \wedge 1)$. By the Laplace principle and the above remark we get

$$\lim_{j\to\infty} \lim_{\epsilon\to 0} -\epsilon \log \mathbb{E}\left[e^{-\frac{h_j(X^\epsilon)}{\epsilon}}\right] = \lim_{j\to\infty} \inf_{x\in\mathcal{X}} \{h_j(x) + I(x)\} = I(\xi)$$

and

$$\lim_{j\to\infty} \lim_{\epsilon\to 0} -\epsilon \log \mathbb{E}\left[e^{-\frac{h_j(X^\epsilon)}{\epsilon}}\right] = \lim_{j\to\infty} \inf_{x\in\mathcal{X}} \{h_j(x) + J(x)\} = J(\xi).$$

Thus $I(\xi) = J(\xi)$. □

1.3 Relative Entropy

In this section we explore the concept of relative entropy and its connection with the evaluation of certain integrals through a variational formula.

Let $(E,\mathcal{E})$ be a measurable space and $\mathcal{P}(E)$ be the set of probability measures on E.

Definition 1.4. *For any fixed $\theta \in \mathcal{P}(E)$, the* ***relative entropy*** *is a map $R(\cdot\|\theta)$ from $\mathcal{P}(E)$ into $[0,+\infty]$ defined by*

$$R(\gamma\|\theta) \doteq \int_E \left(\log \frac{d\gamma}{d\theta}\right) d\gamma,$$

whenever $\gamma \in \mathcal{P}(E)$ is absolutely continuous with respect to θ, and $R(\gamma\|\theta) \doteq +\infty$ otherwise. Clearly $R(\cdot\|\cdot) : \mathcal{P}(E) \times \mathcal{P}(E) \to [0,+\infty]$.

Let us remark that the above definition is actually well-posed. Indeed, if γ is absolutely continuous with respect to θ the Radon-Nikodym derivative $d\gamma/d\theta$ exists in L^1. Moreover, observe that the function $s(\log s)^-$ is bounded for $s \in [0, \infty)$. Thus,

$$\int_E \left(\log \frac{d\gamma}{d\theta}\right)^- d\gamma = \int_E \frac{d\gamma}{d\theta} \left(\log \frac{d\gamma}{d\theta}\right)^- d\theta < \infty.$$

It follows that the integral makes sense and that

$$\int_E \left(\log \frac{d\gamma}{d\theta}\right) d\gamma = \int_E \frac{d\gamma}{d\theta} \left(\log \frac{d\gamma}{d\theta}\right) d\theta$$

We have introduced the relative entropy in order to prove a variational formula which is the starting point for deriving large deviation principles through a weak convergence approach.

Proposition 1.1. *Let $(E, \mathcal{E})$ be a measurable space, k a bounded $\mathcal{E}-$measurable function which maps E into $\mathbb{R}$, and θ a probability measure on $(E, \mathcal{E})$. Then the following variational formula holds*

$$-\log \int_E e^{-k}\, d\theta = \inf_{\gamma \in \mathcal{P}(E)} \left\{ R(\gamma \| \theta) + \int_E k\, d\gamma \right\} \tag{1.7}$$

Moreover, let γ_0 denote the probability measure on E which is absolutely continuous with respect to θ and has Radon-Nikodym derivative

$$\frac{d\gamma_0}{d\theta}(x) \doteq \frac{e^{-k(x)}}{\int_E e^{-k}\, d\theta}$$

then, the infimum in (1.7) *is attained uniquely at γ_0.*

Before demonstrating the proposition, we state and prove two properties about the relative entropy which we shall use in a moment.

Lemma 1.3. *Let $(E, \mathcal{E})$ be a measurable space and $\gamma, \theta \in \mathcal{P}(E)$ be two probability measures on E. Then $R(\gamma\|\theta) \geq 0$ and $R(\gamma\|\theta) = 0$ if and only if $\gamma = \theta$.*

Proof. Clearly, if $R(\gamma\|\theta) = +\infty$ there is nothing to prove, so we can assume that $R(\gamma\|\theta) < +\infty$ and hence that γ is absolutely continuous with respect to θ. Next recall that $s \log s \geq s - 1$ for $s \geq 0$ with equality if and only if $s = 1$. Thus

$$R(\gamma\|\theta) = \int_E \frac{d\gamma}{d\theta} \log \frac{d\gamma}{d\theta}\, d\theta \geq \int_E \frac{d\gamma}{d\theta} - 1\, d\theta = 0$$

with equality if and only if $d\gamma/d\theta = 1$ θ-a.e. if and only if $\gamma = \theta$. □

We are now ready to prove the variational formula.

Proof of Proposition 1.1. It is clear that it suffices to prove (1.7) over the set of probability measures such that $R(\gamma\|\theta) < +\infty$. Thus, we show

$$-\log\int_E e^{-k}\theta = \inf\left\{R(\gamma\|\theta) + \int_E k\,d\gamma : \gamma\in\mathcal{P}(E), R(\gamma\|\theta) < \infty\right\}$$

We have already remarked in the preceding proof that, if $R(\gamma\|\theta) < \infty$, then γ is absolutely continuous with respect to θ. Let γ be such that $R(\gamma\|\theta) < +\infty$. Since θ is absolutely continuous with respect to γ_0, then γ is also absolutely continuous with respect to γ_0. Therefore, we have

$$\begin{aligned}
R(\gamma\|\theta) + \int_E k\,d\gamma &= \int_E \left(\log\frac{d\gamma}{d\theta}\right)d\gamma + \int_E k\,d\gamma \\
&= \int_E \left(\log\frac{d\gamma}{d\gamma_0}\right)d\gamma + \int_E \left(\log\frac{d\gamma_0}{d\theta}\right)d\gamma + \int_E k\,d\gamma \\
&= R(\gamma\|\gamma_0) + \int_E \left(\log\frac{e^{-k}}{\int_E e^{-k}\,d\theta}\right)d\gamma + \int_E k\,d\gamma \\
&= R(\gamma\|\gamma_0) - \log\int_E e^{-k}\,d\theta
\end{aligned}$$

It follows by the previous lemma that

$$R(\gamma\|\theta) + \int_E k\,d\gamma \geq -\log\int_E e^{-k}\,d\theta,$$

with equality if and only if $R(\gamma\|\gamma_0) = 0$, and again because of the lemma, if and only if $\gamma_0 = \gamma$. □

2
Representation formula

2.1 Representation formula for Brownian motion

In this section we state a variational representation formula for the Brownian motion. The formula is due to Dupuis and Bou and the proof can be found in Bou and Dupuis [1997].

We are interested in this formula since it becomes useful in deriving various results in the field of large deviation theory.

Let W be a standard d-dimensional Brownian motion. Then for any measurable and bounded function f mapping $C([0,T];\mathbb{R}^d)$ into $\mathbb{R}$, the following holds

$$-\log \mathbb{E}\left[e^{-f(W)}\right] = \inf_v \mathbb{E}\left[\frac{1}{2}\int_0^T \|v_s\|^2\,ds + f\left(W + \int_0^{\cdot} v_s\,ds\right)\right],$$

where $\mathbb{E}$ denotes the expectation with respect to the probability space where the Brownian motion is defined, and the infimum is over all processes that are progressively measurable with respect to the augmentation of the filtration generated by the Brownian motion.

The importance of the previous representation resides in the fact that we can characterize exponential functionals whenever the functional of interest can be expressed in terms of a measurable functional of the Brownian motion. In this work we shall see an application to the integrals of small noise Itô processes.

From now on we shall restrict our attention to the canonical probability space $(\mathcal{W}^m, \mathcal{B}, \theta)$, where $\mathcal{W}^m = C([0,T];\mathbb{R}^m)$ is a Polish space under the supremum norm, $\mathcal{B} = \mathcal{B}(C([0,T];\mathbb{R}^m))$ is the Borel σ-algebra, θ is the m-dimentional Wiener measure, and $T > 0$ is a fixed real number. Under θ the process $W : [0,T] \times \mathcal{W}^m \to \mathbb{R}^m$ defined by

$$W_t(\omega) \doteq \omega(t) \qquad (t,\omega) \in [0,T] \times \mathcal{W}^m$$

with the filtration $\mathcal{B}_t^W = \sigma(W_s; 0 \leq s \leq t)$, is an m-dimensional Brownian motion. Instead of the natural filtration we shall take into account the augmented filtration $(\mathcal{G}_t)_t$, obtained by considering the collection of sets with measure zero $\mathcal{N}$ and defining for all $t \in [0,T]$

$$\mathcal{G}_t \doteq \sigma(\mathcal{B}_t^W \cup \mathcal{N}).$$

It is a basic property that W_t remains a Brownian motion with respect to the augmented filtration.

Next we recall the definition of progressively measurable processes.

Definition 2.1. *A stochastic process X on $(\mathcal{W}^m, \mathcal{B})$ is said to be **progressively measurable** with respect to the filtration $(\mathcal{G}_t)_t$, if for every t the map*

$$X_s(\omega) : [0,t] \times \mathcal{W}^m \to \mathbb{R}^d$$

is $(\mathcal{B}([0,t]) \otimes \mathcal{G}_t) \setminus \mathcal{B}(\mathbb{R}^d)$-measurable.

Definition 2.2. *We denote by $\mathcal{M}^2[0,T]$ the set of all $\mathcal{G}_t$-progressively measurable processes $v : [0,T] \times \mathcal{W} \to \mathbb{R}^d$ satisfying*

$$\mathbb{E}\left[\int_0^T \|v_s\|^2 \, ds\right] < \infty$$

where $\|\cdot\|$ denotes the standard norm in $\mathbb{R}^d$. We say that $v \in \mathcal{M}_b^2[0,T]$ provided that there exists $C > 0$ such that

$$\int_0^T \|v_s(\omega)\|^2 \, ds < C,$$

for θ-almost all $\omega \in \mathcal{W}^m$.

Clearly, $\mathcal{M}_b^2[0,T] \subset \mathcal{M}^2[0,T]$. We consider that subset because a lot of technicalities simplify in the proofs which follow.

Now we have all the elements to state precisely the Theorem we have anticipated at the beginning of this chapter.

Theorem 2.1. *Let f be a bounded Borel-measurable function which maps $\mathcal{W}^m$ into $\mathbb{R}$. Then*

$$-\log \mathbb{E}\left[e^{-f(W)}\right] = \inf_{v \in \mathcal{M}^2[0,T]} \mathbb{E}\left[\int_0^T \|v_s\|^2 \, ds + f\left(W + \int_0^{\cdot} v_s \, ds\right)\right]. \quad (2.1)$$

Proof. For the proof we refer to Bou and Dupuis [1997]. Here we provide only a sketch of the main ideas for the proof of the **upper bound**, to stress the importance of the variational formula given in Proposition 1.1.

Consider $v \in \mathcal{M}_b^2[0,T]$, then by Girsanov's theorem the process

$$\tilde{W}_t = W_t - \int_0^t v_s\, ds, \qquad t \in [0,T] \tag{2.2}$$

is a d-dimensional Brownian motion with respect to a measure γ_v whose Radon-Nykodim derivative with respect to θ is given by

$$\frac{d\gamma_v}{d\theta} = \exp\left[\int_0^T v_s\, dW_s - \frac{1}{2}\int_0^T \|v_s\|^2\, ds\right]. \tag{2.3}$$

Using the definition of $R(\gamma_v\|\theta)$ and (2.3) we get

$$R(\gamma_v\|\theta) = \int_{\mathcal{W}^d} \log\frac{d\gamma_v}{d\theta}\, d\gamma_v = \mathbb{E}^v\left[\int_0^T v_s\, dW_s - \frac{1}{2}\int_0^T \|v_s\|^2\, ds\right].$$

where $\mathbb{E}^v$ denotes the expectation with respect to γ_v. Thanks to (2.2) and to the martingale property of the stochastic integral,

$$\begin{aligned} R(\gamma_v\|\theta) =& \mathbb{E}^v\left[\int_0^T v_s\, d\tilde{W}_s + \int_0^T \|v_s\|^2\, ds - \frac{1}{2}\int_0^T \|v_s\|^2\, ds\right] \\ =& \mathbb{E}^v\left[\frac{1}{2}\int_0^T \|v_s\|^2\, ds\right]. \end{aligned}$$

Consequently,

$$R(\gamma_v\|\theta) + \int f\, d\gamma_v = \mathbb{E}^v\left[\frac{1}{2}\int_0^T \|v_s\|^2\, ds + f\left(\tilde{W} + \int_0^{\cdot} v_s\, ds\right)\right].$$

Now from Proposition 1.1 we obtain

$$-\log\mathbb{E}\left[e^{-f(W)}\right] \leq \inf_{v\in\mathcal{M}_b^2[0,T]} \mathbb{E}^v\left[\int_0^T \|v_s\|^2\, ds + f\left(\tilde{W} + \int_0^{\cdot} v_s\, ds\right)\right].$$

This is not the desired upper bound, since the expectation depends on v. Using the preceding formula and through an approximation argument it is possible to find

$$-\log\mathbb{E}\left[e^{-f(W)}\right] \leq \mathbb{E}\left[\int_0^T \|v_s\|^2\, ds + f\left(W + \int_0^{\cdot} v_s\, ds\right)\right]$$

for all $v \in \mathcal{M}^2[0,T]$. In the approximation argument it is necessary to exploit some properties of $R(\cdot\|\cdot)$, for example the lower semicontinuity is needed. □

2.2 Representation formula for Itô processes

Let $b(\cdot,\cdot)$ and $\sigma(\cdot,\cdot)$ be predictable functions from $[0,T]\times\mathcal{W}^d$ to $\mathbb{R}^d$ and to $\mathbb{R}^{d\times m}$, respectively (see definition A.1 in the Appendix). Fix $x\in\mathbb{R}^d$, and consider the stochastic functional differential equation

$$dX_t = b(t, X_\cdot)\,dt + \sigma(t, X_\cdot)\,dW_t \tag{2.4}$$

for $t\in[0,T]$ and with initial condition $X_0 = x$.

Suppose that the equation (2.4) has a strong solution. Then there exists a $\mathcal{B}(\mathcal{C}[0,T];\mathbb{R}^m))\setminus\mathcal{B}(\mathcal{C}[0,T];\mathbb{R}^d))$-measurable function

$$h : \mathcal{W}^m \to \mathcal{W}^d$$

such that

$$X = h[W] \qquad \theta\text{-a.e.}$$

[see Rogers and Williams, 2000, Theorem 10.4, page 126]. Hence, for any $f:\mathcal{W}^d\to\mathbb{R}$ bounded and measurable, $f\circ h$ is a bounded and measurable map from $\mathcal{W}^m$ into $\mathbb{R}$. Therefore, the representation formula for the Brownian motion (2.1) implies

$$\begin{aligned} -\log\mathbb{E}\left[e^{-f(X)}\right] &= -\log\mathbb{E}\left[e^{-f\circ h(W)}\right] \\ &= \inf_{v\in\mathcal{M}^2[0,T]}\mathbb{E}\left[\frac{1}{2}\int_0^T\|v_s\|^2\,ds + f\circ h\left(W+\int_0^\cdot v_s\,ds\right)\right] \end{aligned} \tag{2.5}$$

where $\mathbb{E}$ is the expectation with respect to θ.

There are some situations in which we can prove something more. Consider the following controlled stochastic functional differential equation, which is the controlled counterpart of (2.4).

$$dX_t^v = b(t, X_\cdot^v)\,dt + \sigma(t, X_\cdot^v)v_t\,dt + \sigma(t, X_\cdot^v)\,dW_t, \tag{2.6}$$

for $t\in[0,T]$, and with initial condition $X_0^v = x$, where $v:[0,T]\times\mathcal{W}^m\to\mathbb{R}^m$ is a given progressively measurable process, adapted to the augmented filtration generated by W_t, and such that

$$\mathbb{E}\left[\int_0^T\|v_s\|^2\,ds\right] < +\infty$$

If both equations (2.4) and (2.6) have a unique strong solution then the representation formula above simplifies. The next theorem gives a representation formula for diffusions that are strong solutions of (2.4) in terms of the solutions of the controlled equation (2.6).

Theorem 2.2. *Assume that strong uniqueness and strong existence hold for equations* (2.4) *and* (2.6). *Let X be the unique strong solution of equation* (2.4) *and X^v the unique strong solution of* (2.6). *Then for any bounded Borel measurable function $f : \mathcal{C}([0,T];\mathbb{R}^d) \to \mathbb{R}$ the following formula holds*

$$-\log \mathbb{E}\left[e^{-f(X)}\right] = \inf_{v \in \mathcal{M}^2[0,T]} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v_s\|^2\, ds + f(X^v)\right]. \tag{2.7}$$

Proof. By assumption X is a strong solution to (2.4). Therefore, by the previous discussion the representation formula for the Brownian motion yields

$$\begin{aligned} -\log \mathbb{E}\left[e^{-f(X)}\right] &= -\log \mathbb{E}\left[e^{-f\circ h(W)}\right] \\ &= \inf_{v \in \mathcal{M}^2[0,T]} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v_s\|^2\, ds + f \circ h\left(W + \int_0^{\cdot} v_s\, ds\right)\right]. \end{aligned}$$

The representation (2.7) will be established once we verify that $h[W + \int_0^{\cdot} v_s\, ds] = X^v$ θ-a.e. Such an equality comes from strong existence and uniqueness of solutions of the controlled problem.

For continuous and adapted processes $\tilde{W}$ we consider the map Ψ

$$\Psi(\tilde{W})(\omega) \doteq x + \int_0^{\cdot} b(s, h[\tilde{W}(\omega)]_{\cdot})ds + \left(\int_0^{\cdot} \sigma(s, h[\tilde{W}]_{\cdot})d\tilde{W}_s\right)(\omega),$$

where $\omega \in \mathcal{W}$. The map Ψ is certainly well defined when $\tilde{W}$ is given by

$$\tilde{W}_t(\omega) = \omega(t) + \int_0^t v_s(\omega)ds$$

with $v \in \mathcal{M}^2[0,T]$. In this situation, for θ-almost all $\omega \in \mathcal{W}$,

$$\begin{aligned} \Psi(\tilde{W})(\omega) &= x + \int_0^{\cdot} b(s, h[\tilde{W}(\omega)]_s)ds + \\ &\quad + \int_0^{\cdot} \sigma(s, h[\tilde{W}(\omega)]_s)v_s(\omega)ds + \left(\int_0^{\cdot} \sigma(s, h[\tilde{W}]_s)dW_s\right)(\omega), \end{aligned} \tag{2.8}$$

where W is the coordinate process on $\mathcal{W}^m$. Since $h[W]$ is a solution of (2.4), by construction we have

$$h[W(\omega)] = \Psi(W)(\omega) \qquad \theta\text{a.s..}$$

We claim that $h[\tilde{W}(\omega)] = \Psi(\tilde{W})(\omega)$ θ-almost surely. Notice that if we can prove the claim, then, thanks to (2.8), it will follow

$$\begin{aligned} h[\tilde{W}]_t &= \Psi(\tilde{W})_t \\ &= x + \int_0^t b(s, h[\tilde{W}]_{\cdot})ds + \int_0^t \sigma(s, h[\tilde{W}]_{\cdot})v_s ds + \int_0^t \sigma(s, h[\tilde{W}]_{\cdot})dW_s, \end{aligned}$$

accordingly, $h[\tilde{W}]_t$ is a strong solution of (2.6) with respect to W. By assumption, equation (2.6) has a unique strong solution X^v, therefore, $h[W + \int_0^\cdot v_s\, ds] = X^v$, θ-almost everywhere.

Let us prove our claim. We first derive the formula above when

$$\int_0^T \|v_s(\omega)\|^2\, ds < N \qquad \theta - a.e.$$

By Girsanov's theorem there exists a measure γ equivalent to θ such that $\tilde{W}$ is a m-dimensional Brownian motion with respect to γ. To summarize, we have the canonical set-up

$$(\mathcal{W}^m, \{\mathcal{G}_t\}, \theta, W)$$

and the set up

$$(\mathcal{W}^m, \{\mathcal{G}_t\}, \gamma, \tilde{W}).$$

By Theorem 10.4, page 126 of Rogers and Williams [2000], $h(\tilde{W})_\cdot$ satisfies, on $[0, T]$ and with respect to γ

$$h[\tilde{W}]_\cdot = x + \int_0^\cdot b(s, h[\tilde{W}]_\cdot) ds + \int_0^\cdot \sigma(s, h[\tilde{W}]_\cdot) d\tilde{W}_s,$$

Since γ is equivalent to θ, it follows that

$$h[\tilde{W}]_\cdot = \Psi(\tilde{W})_\cdot \qquad \theta\text{-a.s.}$$

This ends the discussion, at least when $v \in \mathcal{M}_b^2[0, T]$.

The next step is to show that $h[\tilde{W}]_\cdot = \Psi(\tilde{W})_\cdot$ holds even in the case

$$\int_0^T \|v_s(\omega)\|^2\, ds < +\infty \qquad \theta - a.e.$$

We argue through an approximation argument. For any $N \in \mathbb{N}$, define the set

$$\Gamma_N = \left\{ \omega \in \mathcal{W}^m : \int_0^T \|v_s(\omega)\|^2\, ds < N \right\}$$

and the processes $v_s^N(\omega) = v_s(\omega) 1_{\Gamma^N}(\omega)$, $\tilde{W}^N(\omega) = W(\omega) + \int_0^\cdot v_s^N(\omega)\, ds$. We observe that $\Gamma_N \subset \Gamma_{N+1}$ and $\theta(\bigcup_{N \in \mathbb{N}} \Gamma_N) = 1$.

Since each $v^N \in \mathcal{M}_b^2[0, T]$, by the previous result, there exists an event G_N with probability one such that

$$h[\tilde{W}^N(\omega)]_\cdot = \Psi(\tilde{W}^N)(\omega)_\cdot$$

for all $\omega \in G_N$. Let us define $G = \bigcap_{N \in \mathbb{N}} G_N$, clearly $\theta(G) = 1$. Next we notice that for each $N \in \mathbb{N}$ we find $H_N \subset \Gamma_N$ with measure 0 such that for all $\omega \in \Gamma_N \cap H_N^c$

$$\Psi(\tilde{W}^N)(\omega) = \Psi(\tilde{W})(\omega)$$

Indeed, it suffices to notice that $\tilde{W} = \tilde{W}^N$ on Γ_N, hence, by the property of locality of the stochastic integral, there exists $H_N \subset \Gamma_N$ with measure 0 such that for all $\omega \in \Gamma_N \cap H_N^c$ and for all $t \in [0, T]$

$$\int_0^t \sigma(s, h_{\cdot}(\tilde{W}^N))\, dW_s = \int_0^t \sigma(s, h_{\cdot}(\tilde{W}))\, dW_s.$$

Accordingly, for all $\omega \in \Gamma_N \cap H_N^c$ and for all $t \in [0, T]$

$$\begin{aligned}
\Psi(\tilde{W}^N)_t &= \int_0^t b(s, h_{\cdot}(\tilde{W}^N))\, ds + \int_0^t \sigma(s, h_{\cdot}(\tilde{W}^N))\, d\tilde{W}_s^N \\
&= \int_0^t b(s, h_{\cdot}(\tilde{W}))\, ds + \int_0^t \sigma(s, h_{\cdot}(\tilde{W}^N)) v_s^N\, ds + \int_0^t \sigma(s, h_{\cdot}(\tilde{W}^N))\, dW_s \\
&= \int_0^t b(s, h_{\cdot}(\tilde{W}))\, ds + \int_0^t \sigma(s, h_{\cdot}(\tilde{W})) v_s\, ds + \int_0^t \sigma(s, h_{\cdot}(\tilde{W}))\, dW_s \\
&= \int_0^t b(s, h_{\cdot}(\tilde{W}))\, ds + \int_0^t \sigma(s, h_{\cdot}(\tilde{W}))\, d\tilde{W}_s = \Psi(\tilde{W})_t.
\end{aligned}$$

Let us set $H = \bigcup_{N \in \mathbb{N}} H_N$, clearly $\theta(H) = 0$.

We have $\theta(G \cap H^c \cap \bigcup_{N \in \mathbb{N}} \Gamma_N) = 1$. If $\omega \in G \cap H^c \cap \bigcup_{N \in \mathbb{N}} \Gamma_N$, then $\exists M \in \mathbb{N}$ such that $\omega \in G \cap H_M^c \cap \Gamma_M$, therefore,

$$\begin{aligned}
\Psi(\tilde{W})(\omega)_{\cdot} &= \Psi(\tilde{W}^M)(\omega)_{\cdot} \\
&= h[\tilde{W}^M(\omega)]_{\cdot} = h[\tilde{W}(\omega)]_{\cdot}
\end{aligned}$$

where the first equality is due to the fact that $\omega \in H_M^c \cap \Gamma_M$, the second one is true since $\omega \in G \subset G_M$, the last is true since $\omega \in \Gamma_M$. Reading the leftmost term and the rightmost term of the last display, our claim is proved. □

3 LDP for small noise Itô processes

3.1 General LDP

In this section we state and prove a large deviation principle for Itô processes. We make assumptions in order to prove the principle through a weak convergence approach in a very general setting. We shall see in the subsequent sections how to verify the hypotheses in particular situations.

Let us define $\mathcal{W}^d \doteq \mathcal{C}([0,T], \mathbb{R}^d)$ for $d \in \mathbb{N}$. Let, for $\epsilon > 0$, b_ϵ and b be predictable functions mapping $[0,T] \times \mathcal{W}^d$ into $\mathbb{R}^d$, and σ_ϵ and σ predictable maps from $[0,T] \times \mathcal{W}^d$ into $\mathbb{R}^{d \times m}$. Let $(\mathcal{W}^m, \mathcal{B}, \theta)$ be the canonical probability space for a m-dimensional Brownian motion W_t, and let $(\mathcal{G}_t)$ be the augmented filtration generated by W.

Let $x \in \mathbb{R}^d$. For $\epsilon > 0$ we consider the stochastic functional differential equation over $[0,T]$

$$dX_t^\epsilon = b_\epsilon(t, X_\cdot^\epsilon)\, dt + \sqrt{\epsilon}\sigma_\epsilon(t, X_\cdot^\epsilon)\, dW_t, \tag{3.1}$$

and with $v \in \mathcal{M}^2[0,T]$ its controlled counterpart

$$dX_t^{\epsilon,v} = b_\epsilon(t, X_\cdot^{\epsilon,v})\, dt + \sigma_\epsilon(t, X_\cdot^{\epsilon,v}) v_t\, dt + \sqrt{\epsilon}\sigma_\epsilon(t, X_\cdot^{\epsilon,v})\, dW_t, \tag{3.2}$$

both with initial condition $X_0^{\epsilon,v} = X_0^\epsilon = x$. Observe that if $\epsilon = 0$, then the first equation becomes a deterministic functional equation

$$\phi(t) = x + \int_0^t b(s, \phi(\cdot))\, ds.$$

Similarly, if $\epsilon = 0$, and we pick $v \in L^2[0,T]$, then the second equation reduces to

$$\phi(t) = x + \int_0^t b(s, \phi(\cdot))\, ds + \int_0^t \sigma(s, \phi(\cdot)) v_s\, ds. \tag{3.3}$$

Let us introduce the following assumptions:

H1 The coefficients b and σ are predictable and $b(t,\cdot)$, $\sigma(t,\cdot)$ are supposed to be uniformly continuous on compact subsets of $\mathcal{W}^d$, uniformly in $t \in [0,T]$. We require $\sigma(\cdot,\phi) \in L^2[0,T]$ for any $\phi \in \mathcal{W}^d$.

H2 The coefficients b_ϵ, σ_ϵ are predictable maps, $b_\epsilon \to b$ and $\sigma_\epsilon \to \sigma$ uniformly on the whole $\mathcal{W}^d$, uniformly in $t \in [0,T]$.

H3 For all ϵ sufficiently small, we require pathwise uniqueness and existence in the strong sense for the equation (3.1).

H4 For any $v \in L^2([0,T];\mathbb{R}^m)$, the equation (3.3) has a unique solution. In this case we will define the map

$$\Gamma_x : L^2([0,T];\mathbb{R}^m) \longrightarrow \mathcal{W}^d$$

which takes $v \in L^2[0,T]$ to the unique solution of equation (3.3),

H5 We assume that, for all $N \in \mathbb{N}$, the map Γ_x is continuous when restricted to

$$S_N \doteq \left\{ f \in L^2[0,T] : \int_0^T \|f_s\|^2\, ds \le N \right\} \subset L^2[0,T],$$

endowed with the weak topology of L^2.

H6 If $\{\epsilon_n\}$ is a sequence such that $\epsilon_n \to 0$ as $n \to \infty$, and $\{v_n\}_{n\in\mathbb{N}} \subset \mathcal{M}^2[0,T]$ such that there exists a constant $N > 0$ with

$$\sup_{n\in\mathbb{N}} \int_0^T \|v_s^n(\omega)\|^2\, ds < N$$

for θ-almost all $\omega \in \mathcal{W}^m$, then the family $\{X^{\epsilon_n, v_n}\}_{n\in\mathbb{N}}$ is tight and such that

$$\sup_{n\in\mathbb{N}} \int_0^T \mathbb{E}[|\sigma(s, X_\cdot^{\epsilon_n, v_n})|^2]\, ds < +\infty$$

Remark 3.1. *We shall see in the next section that assumption H2 can be weakened. Specifically, we shall require the uniform convergence of b_ϵ and σ_ϵ to b and σ only on the bounded subsets of $\mathcal{W}^d$.*

Remark 3.2. *The hypothesis H5 will play a minor role, since it is needed only to guarantee that the rate function has compact sublevel sets, and accordingly is good.*

Theorem 3.1. *Assume (H1-H6). Then the family $\{X^\epsilon\}$ of solutions of the stochastic differential equation (3.1) with initial condition $X^\epsilon(0) = x$, taking values in the Polish space $\mathcal{W}^d$, satisfies the Laplace principle in $\mathcal{W}^d$ with (good) rate function I_x given by*

$$I_x(\phi) = \inf_{\{v \in L^2([0,T];\mathbb{R}^n):\phi_t = \Gamma_x(v)\}} \frac{1}{2}\int_0^T \|v_t\|^2 \, dt$$

whenever $\{v \in L^2([0,T];\mathbb{R}^n) : \phi_t = \Gamma_x(v)\} \neq \emptyset$, and $I_x(\phi) = \infty$ otherwise.

Proof of the lower bound. The first step in proving the theorem is the Laplace principle lower bound. Fix $x \in \mathbb{R}^d$. We have to show that for any bounded and continuous function f mapping $\mathcal{W}^d$ into $\mathbb{R}$, the following limit relation holds

$$\liminf_{\epsilon \to 0+} -\epsilon \log \mathbb{E}\left[e^{-\frac{f(X^\epsilon)}{\epsilon}}\right] \geq \inf_{\phi \in \mathcal{W}^d} \{f(\phi) + I_x(\phi)\}$$

Clearly it suffices to prove that any sequence $\{\epsilon_n\}_{n\in\mathbb{N}}$ such that $\epsilon_n \to 0$ has a subsequence for which the above limit relation holds.

Let $\{\epsilon_n\}_{n\in\mathbb{N}}$ be such that $\epsilon_n \to 0$. By assumption H3, for any $n \in \mathbb{N}$, $X^n \doteq X^{\epsilon_n}$ is a strong solution of equation (3.1). Therefore, as in Section 2.2, there exists a measurable map $h^n : \mathcal{W}^m \to \mathcal{W}^d$ such that $X^n_\cdot = h^n(W)_\cdot$ θ-almost surely. Specifically, the representation formula (2.5) holds, and we get

$$\begin{aligned} -\epsilon_n \log \mathbb{E}\left[e^{-\frac{f(X^n)}{\epsilon_n}}\right] &= -\epsilon_n \log \mathbb{E}\left[e^{-\frac{f\circ h^n(W)}{\epsilon_n}}\right] \qquad (3.4) \\ &= \epsilon_n \inf_{v \in \mathcal{M}^2[0,T]} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v_s\|^2 \, ds + \frac{1}{\epsilon_n} f \circ h^n\left(W + \int_0^\cdot v_s \, ds\right)\right] \\ &= \inf_{v \in \mathcal{M}^2[0,T]} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v_s\|^2 \, ds + f \circ h^n\left(W + \frac{1}{\sqrt{\epsilon_n}}\int_0^\cdot v_s \, ds\right)\right]. \end{aligned}$$

Fix $\delta > 0$. We claim that for every $n \in \mathbb{N}$ there exists $v^n \in \mathcal{M}^2[0,T]$ such that

$$\sup_{n\in\mathbb{N}} \int_0^T \|v^n_s\|^2 \, ds \leq N,$$

where N is a real number to be specified, and such that

$$\begin{aligned} &-\epsilon_n \log \mathbb{E}\left[e^{-\frac{f(X^n)}{\epsilon_n}}\right] \geq \\ &\qquad \geq \mathbb{E}\left[\frac{1}{2}\int_0^T \|v^n_s\|^2 \, ds + f \circ h^n\left(W + \frac{1}{\sqrt{\epsilon_n}}\int_0^\cdot v^n_s \, ds\right)\right] - \delta. \qquad (3.5) \end{aligned}$$

We prove the claim. The definition of infimum implies that for any $n \in \mathbb{N}$ there exists $u^n \in \mathcal{M}^2[0,T]$ such that

$$\begin{aligned} &-\epsilon_n \log \mathbb{E}\left[e^{-\frac{f(X^n)}{\epsilon_n}}\right] \geq \\ &\qquad \geq \mathbb{E}\left[\frac{1}{2}\int_0^T \|u_s^n\|^2\,ds + f\circ h^n\left(W + \frac{1}{\sqrt{\epsilon_n}}\int_0^{\cdot} u_s^n\,ds\right)\right] - \frac{\delta}{2}. \end{aligned}$$

Observe that if we set $M = \|f\|_\infty$, then

$$\begin{aligned} &\mathbb{E}\left[\frac{1}{2}\int_0^T \|u_s^n\|^2\,ds\right] \leq \\ &\leq \mathbb{E}\left[f\circ h^n\left(W + \frac{1}{\sqrt{\epsilon_n}}\int_0^{\cdot} u_s^n\,ds\right)\right] - \epsilon_n \log \mathbb{E}\left[e^{-\frac{f(X^n)}{\epsilon_n}}\right] + \frac{\delta}{2} \leq 2M + \frac{\delta}{2} \end{aligned}$$

Therefore,

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left[\frac{1}{2}\int_0^T \|u_s^n\|^2\,ds\right] \leq 2M + \frac{\delta}{2} < \infty. \tag{3.6}$$

Now define the stopping time

$$\tau_N^n = \inf\left\{t \in [0,T] : \int_0^t \|u_s^n\|^2\,ds \geq N\right\} \wedge T.$$

The processes $u_s^{n,N} = u_s^n 1_{[0,\tau_N^n]}(s)$ belong to $\mathcal{M}_b^2[0,T]$. By Chebichev's inequality and (3.6)

$$\theta(u^n \neq u^{n,N}) \leq \theta\left(\int_0^T \|u_s^n\|^2\,ds \geq N\right) \leq \frac{2M + \delta/2}{N}.$$

This observation implies that

$$\begin{aligned} -\epsilon_n \log \mathbb{E}\left[e^{-\frac{f(X^n)}{\epsilon_n}}\right] &\geq \mathbb{E}\left[\frac{1}{2}\int_0^T \|u_s^{n,N}\|^2\,ds + f\circ h^n\left(W + \frac{1}{\sqrt{\epsilon_n}}\int_0^{\cdot} u_s^{n,N}\,ds\right)\right] \\ &\quad - \frac{M(4M+\delta)}{N} - \frac{\delta}{2} \end{aligned} \tag{3.7}$$

The computation for (3.7) is not difficult, although we write it down for completeness. It suffices to observe that

$$\begin{aligned} &\mathbb{E}\left[\frac{1}{2}\int_0^T \|u_s^n\|^2\,ds + f\circ h^n\left(W + \frac{1}{\sqrt{\epsilon_n}}\int_0^{\cdot} u_s^n\,ds\right)\right] \\ &\geq E\left[\frac{1}{2}\int_0^T \|u_s^{n,N}\|^2\,ds + f\circ h^n\left(W + \frac{1}{\sqrt{\epsilon_n}}\int_0^{\cdot} u_s^{n,N}\,ds\right)\right] + \\ &\mathbb{E}\left[\left(f\circ h^n\left(W + \frac{1}{\sqrt{\epsilon_n}}\int_0^{\cdot} u_s^n\,ds\right) - f\circ h^n\left(W + \frac{1}{\sqrt{\epsilon_n}}\int_0^{\cdot} u_s^{n,N}\,ds\right)\right)1_{\{u^n\neq u^{n,N}\}}\right] \\ &\geq E\left[\frac{1}{2}\int_0^T \|u_s^{n,N}\|^2\,ds + f\circ h^n\left(W + \frac{1}{\sqrt{\epsilon_n}}\int_0^{\cdot} u_s^{n,N}\,ds\right)\right] - 2M\theta(u^n \neq u^{n,N}) \end{aligned}$$

Using the estimate about the event $\{u^n \neq u^{n,N}\}$ one finds the result. Finally, if in (3.7), we take N such that

$$\frac{M(4M+\delta)}{N} < \frac{\delta}{2}$$

and we set $v^n = u^{n,N}$, then the claim is proved.

Let $\{v^n\} \subset \mathcal{M}^2[0,T]$ as above. If we can prove that the controlled stochastic equation

$$dX_t^{n,v^n} = b_{\epsilon_n}(t, X_\cdot^{n,v^n})\,dt + \sigma_{\epsilon_n}(t, X_\cdot^{n,v^n})v_t^n\,dt + \sqrt{\epsilon_n}\sigma_{\epsilon_n}(t, X_\cdot^{n,v^n})\,dW_t,$$

has a unique strong solution, then we will be able to show that

$$h^n\left(W + \frac{1}{\sqrt{\epsilon_n}}\int_0^\cdot v_s^n\,ds\right) = X^{n,v^n} \qquad \theta\text{-a.e,} \tag{3.8}$$

by arguing as in the proof of Theorem 2.2. Let us define the process

$$\tilde{W}_t \doteq W_t + \frac{1}{\sqrt{\epsilon_n}}\int_0^t v_s^n\,ds$$

Since

$$\int_0^t \|v_s^n\|^2\,ds \leq N \qquad \theta\text{-a.e,}$$

Girsanov's theorem is applicable and accordingly there exists a measure γ over $\mathcal{W}^m$ equivalent to θ such that $\tilde{W}_t$ is a $\mathcal{G}_t$-brownian motion on $[0,T]$ [see, for example Karatzas and Shreve, 1991, page 191, Theorem 5.2]. With respect to the measure γ the controlled equation becomes

$$dX_t^{n,v^n} = b_{\epsilon_n}(t, X_\cdot^{n,v^n})\,dt + \sqrt{\epsilon_n}\sigma_{\epsilon_n}(t, X_\cdot^{n,v^n})\,d\tilde{W}_t. \tag{3.9}$$

We have already seen that there exists a strong solution X^{n,v^n} to the controlled equation in the proof of Theorem 2.2. The uniqueness of solutions to the controlled equation follows easily by assumption H3. Indeed, if X and Y are two solutions of the controlled equation, then they are in particular solutions of equation (3.9) under γ and with respect to $\tilde{W}$. Thus, hypothesis H3 of pathwise uniqueness implies that X, Y are indistinguishable.

This discussion implies that for any $n \in \mathbb{N}$, using (3.8), we can rewrite (3.5) and obtain

$$-\epsilon_n \log \mathbb{E}\left[e^{-\frac{f(X^n)}{\epsilon_n}}\right] \geq \mathbb{E}\left[\frac{1}{2}\int_0^T \|v_s^n\|^2\,ds + f(X^{n,v^n})\right] - \delta$$

where X^{n,v^n} is the unique strong solution to the controlled equation (3.2) with control v^n.

Next we would like to prove that (X^{n,v^n}, v^n) is tight as a family of random variables with values in $\mathcal{W}^d \times S_N$. Recall that to prove that (X^{n,v^n}, v^n) is tight, it suffices to show that $\{X^{n,v^n}\}_{n\in\mathbb{N}}$ is tight as a family of $\mathcal{W}^d$-valued random variables and that $\{v^n\}_{n\in\mathbb{N}}$ is tight as a family of S_N-valued random variables, since both S_N and $\mathcal{W}^d$ are Polish spaces. Tightness of the family $\{X^{n,v^n}\}_{n\in\mathbb{N}}$ follows immediately by assumption H6. Tightness of v^n follows by the fact that we are assuming that for all $n \in \mathbb{N}$, v^n takes values in

$$S_N = \left\{ f \in L^2[0,T] : \int_0^T \|f_s\|^2 \, ds \leq N \right\}$$

and S_N is a compact Polish space when endowed with the weak topology of L^2.

Therefore, possibly taking a subsequence, we have that (X^{n,v^n}, v^n) converges in distribution to a random vector (X, v) defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and with values in $\mathcal{W}^d \times S_N$. Let us denote by $\mathbb{E}_{\mathbb{P}}$ the expectation with respect to the measure $\mathbb{P}$. We now prove that X satisfies the equation

$$X_t = x + \int_0^t b(s, X_\cdot) \, ds + \int_0^t \sigma(s, X_\cdot) v_s \, ds. \qquad \mathbb{P}\text{-a.s.} \tag{3.10}$$

In order to do so, for each $t \geq 0$, we consider the map $\Psi_t : \mathcal{W}^d \times S_N \to \mathbb{R}$ defined by

$$\Psi_t(\phi, v) = \left| \phi(t) - x - \int_0^t b(s, \phi(s)) \, ds - \int_0^t \sigma(s, \phi(s)) v_s \, ds \right| \wedge 1.$$

Clearly, Ψ_t is bounded, furthermore it is continuous. Indeed, let $\phi_n \to \phi$ in $\mathcal{W}^d$ and $v_n \to v$ in S_N with respect to the weak topology of L^2 and compute

$$\begin{aligned} &|\Psi_t(\phi_n, v_n) - \Psi_t(\phi, v)| \leq |\phi_n(t) - \phi(t)| + \int_0^t |b(s, \phi_n) - b(s, \phi)| \, ds \\ &+ \int_0^t |\sigma(s, \phi) - \sigma(s, \phi_n)| |v_n(s)| \, ds + \left| \int_0^t \sigma(s, \phi) v(s) \, ds - \int_0^t \sigma(s, \phi) v_n(s) \, ds \right|. \end{aligned}$$

Notice that the set $\mathcal{C} = \{\phi_n : n \in \mathbb{N}\} \cup \{\phi\}$ is a compact subset of $\mathcal{W}^d$, therefore by assumption H1 there exist moduli of continuity ρ_b and ρ_σ mapping $[0, \infty[$ into $[0, \infty[$ such that $|b(s, \phi_\cdot) - b(s, \psi_\cdot)| \leq \rho_b(\|\phi - \psi\|_\infty)$ and $|\sigma(s, \phi_\cdot) - \sigma(s, \psi_\cdot)| \leq \rho_\sigma(\|\phi - \psi\|_\infty)$ for all $s \in [0, T]$ and for all $\phi, \psi \in \mathcal{C}$. Then we get, using Hölder's inequality and $\|v^n\|_{L^2} \leq \sqrt{N}$,

$$\begin{aligned} &|\Psi_t(\phi_n, v_n) - \Psi_t(\phi, v)| \leq \|\phi_n - \phi\|_\infty + t\rho_b(\|\phi_n - \phi\|_\infty) \\ &+ \sqrt{N t} \rho_\sigma(\|\phi_n - \phi\|_\infty) + \left| \int_0^t \sigma(s, \phi) v(s) \, ds - \int_0^t \sigma(s, \phi) v_n(s) \, ds \right|. \end{aligned}$$

Next observe that for any fixed $\phi \in \mathcal{W}^d$ the function $\sigma(\cdot,\phi)$ is in $L^2[0,T]$ by assumption H1, accordingly, the rightmost term of the previous display goes to 0 as $n \to \infty$, since v^n converges weakly to v. Hence, continuity of Ψ_t is proved.

Since (X^{n,v^n}, v^n) converges in distribution to (X, v) and Ψ_t is bounded and continuous, the continuous mapping theorem implies

$$\lim_{n\to\infty} \mathbb{E}[\Psi_t(X^{n,v^n}, v^n)] = \mathbb{E}_{\mathbb{P}}[\Psi_t(X,v)] \tag{3.11}$$

If we show that the limit is actually zero, then, by definition of Ψ_t, X will satisfy equation (3.10) $\mathbb{P}$-almost surely for all $t \in [0,T]$. Since X has continuous paths, it will follow that X satisfies equation (3.10) for all $t \in [0,T]$, $\mathbb{P}$-almost surely.

Observe that

$$\begin{aligned}\mathbb{E}[\Psi_t(X^{n,v^n}, v^n)] &\leq \mathbb{E}\left[\int_0^t |b_{\epsilon_n}(s, X^{n,v^n}_\cdot) - b(s, X^{n,v^n}_\cdot)|\, ds\right] + \\ &+ \mathbb{E}\left[\int_0^t |\sigma_{\epsilon_n}(s, X^{n,v^n}_\cdot) - \sigma(s, X^{n,v^n}_\cdot)||v^n_s|\, ds\right] + \sqrt{\epsilon_n}\mathbb{E}\left[\left|\int_0^t \sigma_{\epsilon_n}(s, X^{n,v^n}_\cdot)\, dW_s\right|\right]\end{aligned}$$

This is easily seen to go to zero as $n \to \infty$. In fact, using the uniform convergence of σ_ϵ to σ and of b_ϵ to b on the whole $\mathcal{W}^d$ uniformly in $t \in [0,T]$, given by assumption H2, we get

$$\begin{aligned}\mathbb{E}[\Psi_t(X^{n,v^n}, v^n)] &\leq t\|b_{\epsilon_n} - b\|_\infty + \\ &+ \|\sigma_{\epsilon_n} - \sigma\|_\infty \mathbb{E}\left[\int_0^T |v^n_s|\, ds\right] + \sqrt{\epsilon_n}\sqrt{\int_0^t \mathbb{E}\left[|\sigma_{\epsilon_n}(s, X^{n,v^n}_\cdot)|^2\right]\, ds}\end{aligned}$$

The last term goes to zero because of assumption H6. Indeed, for the square rooted addendum we have the estimate

$$\begin{aligned}\int_0^t \mathbb{E}\left[|\sigma_{\epsilon_n}(s, X^{n,v^n}_\cdot)|^2\right]\, ds &\leq 2\int_0^T \mathbb{E}\left[|\sigma(s, X^{n,v^n}_\cdot)|^2\right]\, ds + \\ &+ 2\int_0^T \mathbb{E}\left[|\sigma_{\epsilon_n}(s, X^{n,v^n}_\cdot) - \sigma(s, X^{n,v^n}_\cdot)|^2\right]\, ds \\ &\leq 2T\|\sigma_{\epsilon_n} - \sigma\|^2_\infty + 2\int_0^T \mathbb{E}\left[|\sigma(s, X^{n,v^n}_\cdot)|^2\right]\, ds < K\end{aligned}$$

for all $n \in \mathbb{N}$ and some positive constant $K > 0$. Thus, when $\sqrt{\epsilon_n} \to 0$, the last term goes to zero.

Thanks to the preceding discussion and to (3.11), we have proved that

$$\lim_{n\to\infty} \mathbb{E}[\Psi_t(X^{n,v^n}, v^n)] = \mathbb{E}_{\mathbb{P}}[\Psi_t(X,v)] = 0.$$

Thus, X satisfies equation (3.10) for all $t \in [0,T]$, $\mathbb{P}$-almost surely.

Before going further with the computations, let us make a remark. Since $\{v^n\}$ is a sequence of elements such that $\sup_{n\in\mathbb{N}} \int_0^T \|v_s^n\|^2\, ds \le N$ and since we have that v^n converges in distribution to v with respect to the weak topology of L^2, it holds that

$$\liminf_{n\to\infty} \mathbb{E}\left[\int_0^T \|v_s^n\|^2\, ds\right] \ge \mathbb{E}_{\mathbb{P}}\left[\int_0^T \|v_s\|^2\, ds\right].$$

This is due to the fact that the mapping from S_N with the weak topology into $\mathbb{R}$ which takes

$$S_N \ni f \to \int_0^T \|f_s\|^2\, ds \in \mathbb{R}$$

is nonnegative and lower semicontinuous, hence by a version of Fatou's lemma, namely, Theorem A.3.12 page 307 of Dupuis and Ellis [1997] the above relation holds.

Now we have all the ingredients to prove the Laplace principle lower bound. Since (X^{n,v^n}, v^n) converges in distribution to (X, v) and f is bounded and continuous, the continuous mapping theorem and the above remark imply

$$\begin{aligned}
\liminf_{n\to\infty} & -\epsilon_n \log \mathbb{E}\left[e^{-\frac{f(X^n)}{\epsilon_n}}\right] \\
&\ge \liminf_{n\to\infty} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v_s^n\|^2\, ds + f(X^{n,v^n})\right] - \delta \\
&\ge \liminf_{n\to\infty} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v_s^n\|^2\, ds\right] + \lim_{n\to\infty} \mathbb{E}\left[f(X^{n,v^n})\right] - \delta \\
&\ge \mathbb{E}_{\mathbb{P}}\left[\frac{1}{2}\int_0^T \|v_s\|^2\, ds + f(X^v)\right] - \delta \\
&\ge \inf_{\{(v,\phi)\in L^2\times\mathcal{W}^d : \phi_t = \Gamma_x(v)\}} \left\{\frac{1}{2}\int_0^T \|v_s\|^2\, ds + f(\phi)\right\} - \delta \\
&\ge \inf_{\phi\in\mathcal{W}^d} \{I_x(\phi) + f(\phi)\} - \delta.
\end{aligned}$$

The fourth inequality is obtained by evaluating the random variable term by term.

Since δ has been fixed arbitrarily, the lower bound follows. □

Proof of the upper bound. We now prove the Laplace principle upper bound,

$$\limsup_{\epsilon\to 0} -\epsilon \log \mathbb{E}\left[e^{-\frac{f(X^\epsilon)}{\epsilon}}\right] \le \inf_{\phi\in\mathcal{W}} \{I_x(\phi) + f(\phi)\}.$$

Also in this case it suffices to prove that any sequence $\{\epsilon_n\}_{n\in\mathbb{N}}$ such that $\epsilon_n \to 0$ has a subsequence for which the above limiting relation holds.

Choose any $\delta > 0$. If the right hand side is not finite the conclusion is trivial, hence there is no problem in assuming it is finite. For any f bounded and continuous there exists $\phi \in \mathcal{W}^d$ such that

$$I_x(\phi) + f(\phi) \leq \inf_{\psi \in \mathcal{W}} \{I_x(\psi) + f(\psi)\} + \frac{\delta}{2} < \infty. \tag{3.12}$$

For such ϕ, choose $\tilde{v} \in L^2[0,T]$ such that

$$\frac{1}{2}\int_0^T \|\tilde{v}_s\|^2\, ds \leq I_x(\phi) + \frac{\delta}{2},$$

and $\phi_t = x + \int_0^t b(s,\phi_s)\, ds + \int_0^t \sigma(s,\phi_s)\tilde{v}_s\, ds$. This choice is always possible by the definition of I_x and since $I_x(\phi) < \infty$; recall that

$$I_x(\phi) = \inf_{\{v \in L^2([0,T];\mathbb{R}^n):\phi_t = \Gamma_x(v)\}} \frac{1}{2}\int_0^T \|v_t\|^2\, dt$$

whenever $\{v \in L^2([0,T];\mathbb{R}^n) : \phi_t = \Gamma_x(v)\} \neq \emptyset$, and $I_x(\phi) = \infty$ otherwise.

For the chosen ϕ, $\tilde{v}$, consider $X^{\epsilon_n,\tilde{v}}$ as in the proof of the lower bound. Then $(X^{\epsilon_n,\tilde{v}}, \tilde{v})$ is tight, therefore, possibly taking a subsequence $(X^{\epsilon_n}, \tilde{v})$ converges in distribution to a random variable $(X, \tilde{v})$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where X solves the (deterministic) integral equation

$$X_t = x + \int_0^t b(s, X_.)\, ds + \int_0^t \sigma(s, X_.)\tilde{v}_s\, ds, \quad t \in [0,T], \mathbb{P}\text{-a.s.}$$

The solution to that equation is unique by assumption $H4$, therefore $X_t = \phi_t$ for all $t \in [0,T]$, $\mathbb{P}$-almost surely. Now, using representation (3.4) and recalling that $\tilde{v} \in L^2$ is deterministic, we get

$$\begin{aligned}
&\limsup_{n\to\infty} -\epsilon_n \log \mathbb{E}\left[e^{-\frac{f(X^{\epsilon_n})}{\epsilon_n}}\right] \\
&\quad = \limsup_{n\to\infty} \inf_{v \in \mathcal{M}^2[0,T]} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v_s\|^2\, ds + f \circ h^n\left(W + \frac{1}{\sqrt{\epsilon_n}}\int_0^\cdot v_s\, ds\right)\right] \\
&\quad \leq \limsup_{n\to\infty} \mathbb{E}\left[\frac{1}{2}\int_0^T \|\tilde{v}_s\|^2\, ds + f(X^{\epsilon_n,\tilde{v}})\right] \\
&\quad = \frac{1}{2}\int_0^T \|\tilde{v}_s\|^2\, ds + \lim_{n\to\infty} \mathbb{E}[f(X^{\epsilon_n,\tilde{v}})] \\
&\quad \leq I_x(\phi) + \frac{\delta}{2} + \lim_{n\to\infty} \mathbb{E}[f(X^{\epsilon_n,\tilde{v}})].
\end{aligned}$$

Since f is bounded and continuous and $X^{\epsilon_n,\tilde{v}}$ converges in distribution to $X = \phi$, we have $\lim_{n\to\infty} \mathbb{E}[f(X^{\epsilon_n,\tilde{v}})] = f(\phi)$. Thanks to (3.12), we can end the chain of inequalities by

$$I_x(\phi) + \frac{\delta}{2} + f(\phi) \leq \inf_{\psi \in \mathcal{W}^d} \{I_x(\psi) + f(\psi)\} + \delta.$$

Since δ is arbitrary, the proof of the Laplace principle upper bound is complete. □

Goodness of the rate function. To prove that I_x is actually a good rate function, it remains to check that I_x has compact sublevel sets. This follows from the compactness of

$$S_N \doteq \left\{ f \in L^2[0,T] : \int_0^T \|f_s\|^2 \, ds \leq N \right\},$$

for each N, and by the continuity on these sets of the map Γ_x which maps v to the unique solution of $\phi_t = x + \int_0^t b(s, \phi_s)\, ds + \int_0^t \sigma(s, \phi_s) v_s \, ds$, according to assumption H5. Indeed

$$\{\phi \in \mathcal{W}^d : I_x(\phi) \leq N\} = \bigcap_{\epsilon > 0} \Gamma_x(S_{N+\epsilon})$$

is the intersection of compact sets, therefore it is compact. □

3.2 Lipschitz continuous coefficients

In this section we shall show that hypotheses (H1-H6) hold in some special but remarkable cases. With the notation of the previous section, assume

A1 b and σ satisfy a sublinear growth condition. Specifically, there exists $M > 0$ such that for all $s \in [0,T]$, all $\phi \in \mathcal{W}^d$,

$$|b(s, \phi_\cdot)| \leq M(1 + \sup_{0 \leq u \leq s} |\phi_u|) \qquad |\sigma(s, \phi_\cdot)| \leq M(1 + \sup_{0 \leq u \leq s} |\phi_u|).$$

A2 b and σ are locally Lipschitz continuous. Specifically, for any $R > 0$ there exists $L_R > 0$ such that

$$|b(s, \phi_\cdot) - b(s, \psi_\cdot)| \leq L_R \sup_{0 \leq u \leq s} |\phi_u - \psi_u|$$
$$|\sigma(s, \phi_\cdot) - \sigma(s, \psi_\cdot)| \leq L_R \sup_{0 \leq u \leq s} |\phi_u - \psi_u|$$

whenever $\sup_{0 \leq u \leq s} |\phi_u \vee \psi_u| \leq R$ and $s \in [0,T]$.

A3 b_ϵ and σ_ϵ enjoy both property A1 with a common constant $\overline{M}$ and property A2.

A4 $b_\epsilon, \sigma_\epsilon$ converge as $\epsilon \to 0+$ to b and σ respectively, uniformly on bounded subsets of $\mathcal{W}^d$ uniformly in $s \in [0, T]$.

Remark 3.3. *We distinguish between hypotheses A1-A2 and A3 because A3 is not needed to verify H4 and H5. Observe that A4 is not exactly H2, indeed the convergence is not on the whole $\mathcal{W}^d$, but on the bounded subsets of $\mathcal{W}^d$. We shall see in a while how to deal with this change.*

Remark 3.4. *Assumption A2 implies that if $0 \le t \le T$ and $\phi, \psi \in \mathcal{W}^d$ are such that $\phi_s = \psi_s$ for all $s \in [0, t]$, then $b(t, \phi_.) = b(t, \psi_.)$. Indeed if $R > 0$ is such that $\sup_{0 \le s \le t} |\phi_s \vee \psi_s| \le R$ then*

$$|b(t, \phi_.) - b(t, \psi_.)| \le L_R \sup_{0 \le s \le t} |\phi_s - \psi_s| = 0$$

In particular $b(t, \phi_.) = b(t, \phi_{t \wedge .})$ for all $\phi \in \mathcal{W}^d$.

Theorem 3.2. *Assume (A1-A4). Then the family $\{X^\epsilon\}$ of solutions of the stochastic differential equation (3.1), taking values in the Polish space $\mathcal{W}^d$, satisfies the Laplace principle in $\mathcal{W}^d$ with (good) rate function I_x, given by*

$$I_x(\phi) = \inf_{\{v \in L^2([0,T];\mathbb{R}^n) : \phi_t = \Gamma_x(v)\}} \frac{1}{2} \int_0^T \|v_t\|^2 \, dt$$

whenever $\{v \in L^2([0, T]; \mathbb{R}^n) : \phi_t = \Gamma_x(v)\} \neq \emptyset$, and $I_x(\phi) = \infty$ otherwise.

To prove the large deviation principle we shall now verify hypotheses H1-H6. As we have previously remarked we will not be able to prove H2. Instead, we are going to prove that in this special setting $H2$ is not really needed, this discussion is postponed to the end of the section.

Hypotheses H1, H3

H1 is satisfied, in fact $b(t, \cdot)$ and $\sigma(t, \cdot)$ are uniformly continuous on bounded subsets of $\mathcal{W}^d$, uniformly in $t \in [0, T]$ because of assumption A2. Moreover $\sigma(\cdot, \phi)$ belongs to $L^2[0, T]$ for any $\phi \in \mathcal{W}^d$ because

$$|\sigma(t, \phi)|^2 \le 2M^2(1 + \|\phi\|_\infty^2) < \infty.$$

Finally, by theorem A.1, it follows that assumption H3 of pathwise uniqueness and of strong existence of solutions of equation (3.1) holds.

Hypotheses H4, H5

Define the map $S_x : \mathcal{W}^d \to \mathcal{W}^d$ which takes $\phi \in \mathcal{W}^d$ to

$$S_x(\phi) = x + \int_0^\cdot b(s, \phi(\cdot)) \, ds + \int_0^\cdot \sigma(s, \phi(\cdot)) v_s \, ds,$$

for $x \in \mathbb{R}^d$ and $v \in L^2[0, T]$ fixed.

Lemma 3.1 (Existence). *Let us assume* $A1$ *and* $H1$. *Then, for any choice of* $v \in L^2[0,T]$ *the equation*

$$\phi(t) = S_x(\phi)_t = x + \int_0^t b(s,\phi(\cdot))\,ds + \int_0^t \sigma(s,\phi(\cdot))v_s\,ds. \tag{3.13}$$

has a solution in $\mathcal{W}^d$.

Before proving the theorem, we stress that for existence of solutions to equation (3.13), assumption $A2$ has been weakened. Indeed we only require that $b(t,\cdot)$ and $\sigma(t,\cdot)$ are uniformly continuous on bounded subsets of $\mathcal{W}^d$. Clearly, assumption A1 together with assumption A2 imply H1.

The main idea of the proof is to apply a fixed point theorem, namely that due to Schauder, which we recall next.

Theorem 3.3 (Schauder fixed point theorem). *Let* $\mathcal{X}$ *be a closed and convex subset of a Banach space. Let* $T : \mathcal{X} \to \mathcal{X}$ *be a continuous mapping. If* $T(\mathcal{X})$ *has compact closure, then there exists* $x \in \mathcal{X}$ *such that* $T(x) = x$.

Proof. For a proof see, for instance, Gilbarg and Trudinger [2001, page 279, section 11.1]. □

Proof of Lemma. It is well known that $\mathcal{W}^d$ is a Banach space with respect to the uniform norm $\|\cdot\|_\infty$. In order to apply Schauder fixed point theorem, with the choice $T = S_x$, we shall restrict our attention to a closed and convex subset of $\mathcal{W}^d$. With this in mind, let $g : [0,T] \to \mathbb{R}$ be the map defined by

$$g(t) \doteq \lambda e^{\lambda t} \quad \forall t \in [0,T].$$

where λ is a positive real number to be specified below. Next define

$$\mathcal{X} \doteq \{\phi \in \mathcal{W}^d : \sup_{0\le s\le t} |\phi_u|^2 \le g(t) \text{ for all } t \in [0,T], \phi_0 = x\}.$$

Clearly if $\lambda > |x|^2$ then $\mathcal{X} \neq \emptyset$ since $\phi_\cdot \equiv x \in \mathcal{X}$.

We have to show that:

1. $\mathcal{X}$ is closed and convex;

2. $S_x(\mathcal{X}) \subseteq \mathcal{X}$;

3. $S_x(\mathcal{X})$ is precompact;

4. S_x is continuous.

Step 1. It is an easy computation to check that $\mathcal{X}$ is closed and convex.

$\mathcal{X}$ is closed. Take $\{\phi^n\} \subset \mathcal{X}$ such that $\phi^n \to \phi$, then $\phi \in \mathcal{X}$. In fact it is obvious that $\phi_0 = x$, because uniform convergence implies pointwise convergence, and

$$\sup_{0\le s\le t} |\phi_s| \le \|\phi_n - \phi\|_\infty + \sup_{0\le s\le t} |\phi^n_s| \le \|\phi_n - \phi\|_\infty + \sqrt{g(t)},$$

sending $n \to +\infty$ we get $\sup_{0\le s\le t} |\phi_u| \le \sqrt{g(t)}$, hence $\phi \in \mathcal{X}$.

$\mathcal{X}$ is convex. Let $\mu, \lambda \ge 0$ such that $\mu + \lambda = 1$. Let $\phi, \psi \in \mathcal{X}$ and consider $\mu\phi + \lambda\psi$. Again, it is obvious that $\mu\phi_0 + \lambda\psi_0 = x$. Moreover,

$$\begin{aligned}\sup_{0\le s\le t} |\mu\phi_s + \lambda\psi_s| &\le \mu \sup_{0\le s\le t} |\phi_s| + \lambda \sup_{0\le s\le t} |\psi_s| \\ &\le \mu\sqrt{g(t)} + \lambda\sqrt{g(t)} = \sqrt{g(t)}.\end{aligned}$$

Step 2. Let $\phi \in \mathcal{X}$; the definition of S_x implies immediately that $S_x(\phi)_0 = x$. Next the usual computations involving the sublinear growth at infinity, Hölder's inequality and the elementary inequality $(a+b+c)^2 \le 3a^2 + 3b^2 + 3c^2$, yield

$$\begin{aligned}|S_x(\phi)_t|^2 &\le 3|x|^2 + 3t\int_0^t |b(s,\phi_\cdot)|^2\,ds + 3\|v\|^2_{L^2}\int_0^t |\sigma(s,\phi_\cdot)|^2\,ds \\ &\le 3|x|^2 + 6M^2(t+\|v\|^2)\int_0^t (1 + \sup_{0\le u\le s}|\phi_u|^2)\,ds \\ &\le 3|x|^2 + 6M^2 t(t+\|v\|^2) + 6M^2(t+\|v\|^2)\int_0^t g(s)\,ds.\end{aligned}$$

Therefore,

$$\begin{aligned}\sup_{0\le s\le t} |S_x(\phi)_s|^2 \le 3|x|^2 &+ 6M^2 t(t+\|v\|^2) + \\ &+ 6M^2(t+\|v\|^2)e^{\lambda t} - 6M^2(t+\|v\|^2).\end{aligned}$$

Chose $\lambda > 3|x|^2 + 6M^2T(T+\|v\|^2) + 6M^2(T+\|v\|^2)$. Then, since $e^{\lambda t} \ge 1$ for all $t \ge 0$,

$$\begin{aligned}\sup_{0\le s\le t} |S_x(\phi)_s|^2 &\le \Big(3|x|^2 + 6M^2 t(t+\|v\|^2) + 6M^2(t+\|v\|^2)\Big)e^{\lambda t} \\ &\le \lambda e^{\lambda t} = g(t)\end{aligned}$$

for all $t \in [0,T]$. Hence $S_x(\phi) \in \mathcal{X}$ for all $\phi \in \mathcal{X}$.

Step 3. To prove that $S_x(\mathcal{X})$ has compact closure it suffices to show that $S_x(\mathcal{X})$ is bounded and equicontinuous, so that the conclusion will follow from the Ascoli-Arzel Theorem. Plainly, $S_x(\mathcal{X})$ is bounded since for all $\phi \in S_x(\mathcal{X})$ we have $\|\phi\|_\infty \leq \sqrt{g(T)}$. Thus, it remains to check equicontinuity. To this end, let $\phi \in \mathcal{X}$ and $0 \leq s < t \leq T$, compute

$$\begin{aligned}|S_x(\phi)_t - S_x(\phi)_s| &\leq \int_s^t |b(u,\phi_\cdot)|\, du + \int_s^t |\sigma(u,\phi_\cdot)||v_u|\, du \\ &\leq \int_s^t M(1+\|\phi\|_\infty)\, du + \int_s^t M(1+\|\phi\|_\infty)|v_u|\, du \\ &\leq (t-s)M(1+\sqrt{g(T)}) + M(1+\sqrt{g(T)})\sqrt{t-s}\,\|v\|_{L^2} \\ &\leq \sqrt{t-s}\, M(1+\sqrt{g(T)})(\sqrt{T}+\|v\|_{L^2})\end{aligned}$$

which is exactly the equihölderianity of exponent $1/2$ of the functions in $S_x(\mathcal{X})$.

Step 4. Finally we prove that the map S_x is continuous. Pick $\{\phi^n\}_{n\in\mathbb{N}} \subset \mathcal{X}$ such that $\phi^n \to \phi \in \mathcal{X}$. Observe that $\mathcal{C} = \{\phi_n : n \in \mathbb{N}\} \cup \{\phi\}$ is compact. Then by hypothesis H1 there exist moduli of continuity ρ_b and ρ_σ mapping $[0,\infty[$ into $[0,\infty[$ such that

$$|b(s,\phi_\cdot) - b(s,\psi_\cdot)| \leq \rho_b(\|\phi-\psi\|_\infty) \qquad |\sigma(s,\phi_\cdot) - \sigma(s,\psi_\cdot)| \leq \rho_\sigma(\|\phi-\psi\|_\infty)$$

for all $s \in [0,T]$ and for all $\phi,\psi \in \mathcal{C}$. Thus,

$$\begin{aligned}&|S_x(\phi^n)_t - S_x(\phi)_t| \\ &\quad\leq \int_0^t |b(s,\phi^n_\cdot) - b(s,\phi_\cdot)|\, ds + \int_0^t |\sigma(s,\phi^n_\cdot) - \sigma(s,\phi_\cdot)||v_s|\, ds \\ &\quad\leq T\rho_b(\|\phi^n-\phi\|_\infty) + \sqrt{T}\|v\|_{L^2}\rho_\sigma(\|\phi^n-\phi\|_\infty)\end{aligned}$$

which gives the desired result. □

Lemma 3.2 (Uniqueness). *Let us assume that b and σ satisfy A1 and A2. Then for any $\xi \in \mathbb{R}^d$ and $v \in L^2[0,T]$ there is a unique solution $\phi_\xi : [0,T] \to \mathbb{R}^d$ of equation* (3.13)*, and satisfies, for all $t \in [0,T]$,*

$$\sup_{0\leq s\leq t} |\phi_t|^2 \leq (3|\xi|^2 + 6M^2t^2 + 6M^2t\|v\|^2_{L^2})e^{6M^2t(t+\|v\|^2_{L^2})}. \tag{3.14}$$

Proof. We first prove uniqueness. Let ϕ_ξ and ϕ_η be solutions of equation (3.13) on $[0,T]$ with initial condition $\phi_\xi(0) = \xi$ and $\phi_\eta(0) = \eta$, respectively. We compute

$$\begin{aligned}&|\phi_\xi(t) - \phi_\eta(t)| \leq |\xi-\eta| + \\ &\qquad \int_0^t |b(s,\phi_\xi) - b(s,\phi_\eta)|\, ds + \int_0^t |\sigma(s,\phi_\xi) - \sigma(s,\phi_\eta)||v_s|\, ds.\end{aligned}$$

By taking the square, using Hölder and the local Lipschitz continuity, we get for $R > 0$ big enough

$$|\phi_\xi(t) - \phi_\eta(t)|^2 \leq 3|\xi - \eta|^2 + 3L_R^2(t + \|v\|_{L^2}^2) \int_0^t \sup_{0 \leq u \leq s} |\phi_\xi(u) - \phi_\eta(u)|^2 \, ds$$

Thus, by Gronwall's inequality we get

$$\sup_{0 \leq s \leq t} |\phi_\xi(s) - \phi_\eta(s)|^2 \leq 3|\xi - \eta|^2 e^{3L_R^2(t + \|v\|_{L^2}^2)t},$$

which gives uniqueness, once we take $\xi = \eta$.

Next we show the estimate (3.14):

$$\begin{aligned}
|\phi_t|^2 &\leq 3|\xi|^2 + 3t \int_0^t |b(s, \phi_\cdot)|^2 \, ds + 3 \left(\int_0^t |\sigma(s, \phi_\cdot)||v_s| \, ds \right)^2 \\
&\leq 3|\xi|^2 + 6M^2(t + \|v\|_{L^2}^2) \int_0^t (1 + \sup_{0 \leq u \leq s} |\phi_u|^2) \, ds \\
&\leq (3|\xi|^2 + 6M^2t^2 + 6M^2t\|v\|_{L^2}^2) + 6M^2(t + \|v\|_{L^2}^2) \int_0^t \sup_{0 \leq u \leq s} |\phi_u|^2 \, ds
\end{aligned}$$

Now, Gronwall's inequality yields

$$\sup_{0 \leq s \leq t} |\phi_t|^2 \leq (3|\xi|^2 + 6M^2t^2 + 6M^2t\|v\|_{L^2}^2) e^{6M^2t(t + \|v\|_{L^2}^2)}$$

for all t such that the solution exists. In particular if we have a global solution on $[0, T]$ then it is bounded by a constant which depends increasingly on $\|v\|_{L^2}^2$. □

Lemma 3.3. *Assume A1 and A2, then hypothesis H5 holds. Specifically, for any $N > 0$ the map Γ_x defined in H4 is continuous into $\mathcal{W}^d$ with the supremum norm when restricted to*

$$S_N \doteq \left\{ f \in L^2[0, T] : \int_0^T \|f_s\|^2 \, ds \leq N \right\} \subset L^2[0, T],$$

endowed with the weak topology of L^2.

Proof. We now prove continuity of the map Γ_x. Let $N > 0$ be fixed. Then S_N is a compact Polish space when endowed with the weak topology of L^2. Take $\{v^n\} \subset S_N$ such that $v_n \to v$ weakly, and define $\phi^n = \Gamma_x(v^n)$, $\phi = \Gamma_x(v)$. We want to estimate

$$\|\Gamma_x(v^n) - \Gamma_x(v)\|_\infty = \sup_{t \in [0,T]} |\phi_n(t) - \phi(t)|$$

Since $\|v^n\|_{L^2}^2 \le N$, we have $\sup_{n\in\mathbb{N}} \|\phi^n\|_\infty < +\infty$ by (3.14). Hence, we can apply the assumption of Lipschitz continuity over bounded subsets. Accordingly there exists a common Lipschitz constant $L > 0$ for b and σ.

$$\phi_t^n - \phi_t = \int_0^t (b(s,\phi_\cdot^n) - b(s,\phi_\cdot))\, ds + \\ + \int_0^t (\sigma(s,\phi_\cdot^n) - \sigma(s,\phi_\cdot))v_s^n\, ds + \int_0^t \sigma(s,\phi_\cdot)(v_s^n - v_s)\, ds$$

Hence,

$$\sup_{0\le u\le t} |\phi_u^n - \phi_u| \le L\int_0^t \sup_{0\le u\le s} |\phi_u^n - \phi_u|\, ds + \\ + L\int_0^t \sup_{0\le u\le s} |\phi_u^n - \phi_u||v_s^n|\, ds + \sup_{0\le u\le T} \left|\int_0^u \sigma(s,\phi_\cdot)(v_s^n - v_s)\, ds\right| \quad (3.15)$$

Let us consider for a moment

$$\Delta_\sigma \doteq \sup_{0\le u\le T} \left|\int_0^u \sigma(s,\phi_\cdot)(v_s^n - v_s)\, ds\right|.$$

We prove that Δ_σ goes to zero as $n \to +\infty$. Observe that for any fixed ϕ the function $\sigma(\cdot,\phi_\cdot) \in L^\infty[0,T]$, indeed $|\sigma(s,\phi_\cdot)| \le M(1 + \|\phi\|_\infty)$ for any $s \in [0,T]$. It follows that σv^n converges weakly to σv in L^2. Moreover, the family $\{\sigma v^n\}_{n\in\mathbb{N}}$ is bounded in L^2 with respect to the L^2-norm. Hence

$$\int_0^t \sigma(s,\phi_\cdot)v_s^n\, ds \to \int_0^t \sigma(s,\phi_\cdot)v_s\, ds$$

uniformly for $t \in [0,T]$. In fact, consider a bounded sequence $u_n \in L^2[0,T]$ which converges weakly to $u \in L^2$; define

$$F_n(t) \doteq \int_0^t u_n(s)\, ds \quad \text{and} \quad F(t) \doteq \int_0^t u(s)\, ds.$$

we observe that for $0 \le s \le t \le T$,

$$|F_n(t) - F_n(s)| \le \int_s^t |u_n(r)|\, dr \le \sqrt{t-s}\|u_n\|_{L^2}.$$

Since $\{u_n\}$ is bounded, it follows that $\{F_n\}_{n\in\mathbb{N}}$ is equihölder and accordingly equicontinuous. Moreover, $F_n(0) = 0$, hence we get $|F_n(t)| \le \sqrt{T}\sup_{n\in\mathbb{N}} \|u_n\|_{L^2}$ for every n and every t, that is, the set $\{F_n : n \in \mathbb{N}\}$ is uniformly bounded, so that it has compact closure by Ascoli's theorem; every subsequence has then a uniformly converging subsequence; but pointwise convergence to F, which follows by

$$F_n(t) = \int_0^T 1_{[0,t]}(s)u_n(s)\, ds \to \int_0^T 1_{[0,t]}(s)u(s)\, ds = F(t),$$

implies that all these sequences converge uniformly to F, which is therefore the uniform limit of the entire sequence.

Therefore, for any $\epsilon > 0$ there exists $n_\epsilon \in \mathbb{N}$ such that $\forall n \geq n_\epsilon$

$$\sup_{0\leq u\leq T}\left|\int_0^u \sigma(s,\phi_.)(v_s^n - v_s)\,ds\right| \leq \epsilon$$

Next we take the square on both sides of (3.15), we apply Hölder's inequality and finally through Gronwall's lemma we get

$$\sup_{0\leq u\leq t}|\phi_u^n - \phi_u|^2 \leq 3\epsilon^2 e^{3L^2 t(t+N)}.$$

By taking the supremum over $t \in [0,T]$ we have the desired result

$$\|\phi^n - \phi\|_\infty \leq \sqrt{3}\epsilon e^{\frac{3}{2}L^2 T(T+N)}$$

This ends the proof of continuity. In particular, hypothesis $H5$ holds. □

Hypothesis H6

In this subsection we prove that hypothesis H6 holds.

Let $\{\epsilon_n\}$ be a sequence such that $\epsilon_n \to 0$ as $n \to \infty$, and $\{v_n\}_{n\in\mathbb{N}} \subset \mathcal{M}_b^2[0,T]$ such that there exists a constant $N > 0$ with

$$\sup_{n\in\mathbb{N}}\int_0^T \|v_s^n(\omega)\|^2\,ds < N$$

for θ-almost all $\omega \in \mathcal{W}^m$. We want to prove that the family $\{X^{\epsilon_n,v_n}\}_{n\in\mathbb{N}}$ of solutions to the controlled equation (3.2) is tight.

In order to prove that H6 hold, we need a preliminary result, by which we have an estimate on the solutions.

Lemma 3.4. *Let b and σ be predictable processes which have sub-linear growth at infinity and let $X \in \Lambda^2[0,T]$ satisfy* (2.6)*, for $t \leq T$. Let $p \geq 2$. Then*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|X_t|^p\right] \leq C(T,N,M)(1+|x|^p). \tag{3.16}$$

In particular $X \in \mathcal{M}^p[0,T]$, and the constant $C(T,N,M)$ is increasing in all its arguments.

Proof. The main idea is to estimate the function

$$t \to \mathbb{E}\left[\sup_{0\leq u\leq t}|X_u|^p\right]$$

by Gronwall's lemma, which requires the function to be bounded. Therefore, let τ_R be the exit time of X from the open ball B_R of center 0 and radius R. Since $|X_{t\wedge\tau_R}| \leq R$, then the process $X_{t\wedge\tau_R}$ belongs to $\mathcal{M}^p[0,T]$. Moreover,

$$X_{t\wedge\tau_R} = x + \int_0^t b(s,X_\cdot)1_{\{s\leq\tau_R\}}\,ds + \int_0^t \sigma(s,X_\cdot)v_s 1_{\{s\leq\tau_R\}}\,ds + \\ + \int_0^t \sigma(s,X_\cdot)1_{\{s\leq\tau_R\}}\,dW_s$$

Next we need some preliminary calculations; first observe that

$$|b(s,X_\cdot)1_{\{s\leq\tau_R\}}| \leq M(1+\sup_{0\leq u\leq s}|X_{u\wedge\tau_R}|).$$

Indeed, if $\omega\in\mathcal{W}^m$ is such that $s\leq\tau_R(\omega)$, then

$$|b(s,X_\cdot(\omega))1_{\{s\leq\tau_R\}}(\omega)| = |b(s,X_\cdot(\omega))| \\ \leq M(1+\sup_{0\leq u\leq s}|X_u(\omega)|) \leq M(1+\sup_{0\leq u\leq s}|X_{u\wedge\tau_R}(\omega)|)$$

If $\omega\in\mathcal{W}^m$ is such that $s>\tau_R(\omega)$, then $|b(s,X_\cdot)1_{\{s<\tau_R\}}| = 0$ and the inequality is obviously true. Therefore,

$$I_1 \doteq \mathbb{E}\left[\int_0^t |b(s,X_\cdot)1_{\{s<\tau_R\}}|^p\,ds\right] \leq 2^{p-1}M^p\left(t+\mathbb{E}\left[\int_0^t \sup_{0\leq u\leq s}|X_{u\wedge\tau_R}|^p\,ds\right]\right)$$

For the next calculation, Burkholder-Davis-Gundy's inequality is required [see, for example Karatzas and Shreve, 1991, theorem 3.28 page 166]. Set

$$I_2 \doteq \mathbb{E}\left[\sup_{0\leq u\leq t}\left|\int_0^u \sigma(s,X_\cdot)1_{\{s<\tau_R\}}\,dW_s\right|^p\right] \\ \leq c_p\mathbb{E}\left[\int_0^t |\sigma(s,X_\cdot)1_{\{s<\tau_R\}}|^2\,ds\right]^{\frac{p}{2}} \\ \leq c_p 2^{p-1}t^{\frac{p-2}{2}}M^p\left(t+\mathbb{E}\left[\int_0^t \sup_{0\leq u\leq s}|X_{u\wedge\tau_R}|^p\,ds\right]\right).$$

Finally, using Hölder's inequality, we get

$$I_3 \doteq \mathbb{E}\left[\left(\int_0^t |\sigma(s,X_\cdot)1_{\{s<\tau_R\}}||v_s|\,ds\right)^p\right] \\ \leq \mathbb{E}\left[\left(\int_0^t |\sigma(s,X_\cdot)1_{\{s<\tau_R\}}|^2\,ds\int_0^t |v_s|^2\,ds\right)^{\frac{p}{2}}\right] \\ \leq 2^{p-1}t^{\frac{p-2}{2}}M^pN^{\frac{p}{2}}\left(t+\mathbb{E}\left[\int_0^t \sup_{0\leq u\leq s}|X_{u\wedge\tau_R}|^p\,ds\right]\right).$$

Then we have

$$\mathbb{E}\Big[\sup_{0\le r\le t}|X_{t\wedge\tau_R}|^p\Big] \le 4^{p-1}|x|^p + 4^{p-1}t^{p-1}I_1 + 4^{p-1}I_2 + 4^{p-1}I_3 \tag{3.17}$$
$$\le 4^{p-1}|x|^p + 2^{3p-3}M^p t^{\frac{p-2}{2}}(t^{\frac{p}{2}} + c_p + N^{\frac{p}{2}})\Big(t + \int_0^t \mathbb{E}\Big[\sup_{0\le u\le s}|X_{u\wedge\tau_R}|^p\Big]\,ds\Big)$$

To short our notation we set

$$K(t,N,M) \doteq \max\Big\{4^{p-1}, 2^{3p-3}M^p t^{\frac{p-2}{2}}(t^{\frac{p}{2}} + c_p + N^{\frac{p}{2}}),$$
$$2^{3p-3}M^p t^{\frac{p}{2}}(t^{\frac{p}{2}} + c_p + N^{\frac{p}{2}})\Big\}.$$

Notice that the constant $K(t,N,M)$ is strictly positive and increasing with respect all its arguments. With the new notation (3.17) can be rewritten as

$$\mathbb{E}\Big[\sup_{0\le r\le t}|X_{t\wedge\tau_R}|^p\Big] \le$$
$$\le K(t,N,M)\Big(1 + |x|^p + \int_0^t \mathbb{E}\Big[\sup_{0\le u\le s}|X_{u\wedge\tau_R}|^p\Big]\,ds\Big)$$

We are now in a position to use Gronwall's inequality; thus, we obtain

$$\mathbb{E}\Big[\sup_{0\le r\le t}|X_{r\wedge\tau_R}|^p\Big] \le K(t,N,M)(1+|x|^p)e^{K(t,N,M)t}$$

We observe that the constant $K(t,N,M)$ does not depend on R. Next we notice that

$$\mathbb{E}\Big[\sup_{0\le t\le T\wedge\tau_R}|X_t|^p\Big] = \mathbb{E}\Big[\sup_{0\le t\le T}|X_{t\wedge\tau_R}|^p\Big] \le C(T,N,M)(1+|x|^p)$$

where $C(t,N,M) \doteq K(t,N,M)e^{K(t,N,M)t}$ for simplicity of notation.

The constant $C(t,N,M)$ is strictly positive and is increasing with respect to all its arguments.

Since X is continuous, we have $\sup_{0\le t\le T\wedge\tau_R}|X_t|^p = R^p$ on the event $\{\tau_R \le T\}$. Therefore

$$\mathbb{P}(\tau_R \le T) \le \mathbb{P}\Big(\sup_{0\le t\le T\wedge\tau_R}|X_t|^p \ge R^p\Big)$$
$$\le \frac{1}{R^p}\mathbb{E}\Big[\sup_{0\le t\le T\wedge\tau_R}|X_t|^p\Big] \le \frac{C(T,N,M)(1+|x|^p)}{R^p}$$

Accordingly $\mathbb{P}(\tau_R \leq T) \to 0$ as $R \to +\infty$. Since $\{\tau_{R'} \leq T\} \subset \{\tau_R \leq T\}$ whenever $R' > R$, then $\tau_R \wedge T \to T$ a.e. as $R \to +\infty$. Finally using Fatou's lemma we get

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |X_t|^p\right] \leq \liminf_{R\to+\infty} \mathbb{E}\left[\sup_{0\leq t\leq T\wedge\tau_R} |X_t|^p\right] \leq C(t,N,M)(1+|x|^p)$$

□

Remark 3.5. *In the proof above, we showed in particular that if τ_R is the exit time of X from the ball B_R and $p = 2$, then*

$$\lim_{R\to+\infty} \mathbb{P}(\tau_R \leq T) \leq \lim_{R\to+\infty} \frac{C(T,N,M)(1+|x|^2)}{R^2} = 0$$

The convergence is actually uniform in x belonging to a compact subset of $\mathbb{R}^n$.

Remark 3.6. *Suppose that X^ϵ satisfies*

$$dX^\epsilon_t = b_\epsilon(t, X^\epsilon_\cdot)\, dt + \sigma_\epsilon(t, X^\epsilon_\cdot) v_t\, dt + \sqrt{\epsilon}\sigma_\epsilon(t, X^\epsilon_\cdot)\, dW_t$$

with $X^\epsilon_0 = x$. Let b_ϵ and σ_ϵ have sublinear growth at infinity with constant M. Observe that if $\epsilon < 1$, then $\sqrt{\epsilon}\sigma_\epsilon$ has sublinear growth at infinity with constant M. Hence, the previous computations remain true and we get

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |X^\epsilon_t|^p\right] \leq C(T, \|v\|^2_{L^2}, M)(1+|x|^p). \tag{3.18}$$

Remark 3.7. *In particular, estimate* (3.18) *and sublinear growth at infinity for σ imply*

$$\sup_{n\in\mathbb{N}} \int_0^T \mathbb{E}\left[|\sigma(s, X_\cdot^{\epsilon_n,v_n})|^2\right] ds < +\infty$$

We want to use the previous lemma in order to exploit the tightness property for a family of solutions of the controlled equation (3.2):

$$dX^{\epsilon,v}_t = b_\epsilon(t, X^{\epsilon,v}_\cdot)\, dt + \sigma_\epsilon(t, X^{\epsilon,v}_\cdot) v_t\, dt + \sqrt{\epsilon}\sigma_\epsilon(t, X^{\epsilon,v}_\cdot)\, dW_t, \qquad X^{\epsilon,v}_0 = 0.$$

We assume that there exists a common constant $\overline{M}$ for any $\epsilon > 0$ sufficiently small (which does not depend on ϵ) such that b_ϵ and σ_ϵ have sublinear growth at infinity. This is always true if b_ϵ and σ_ϵ converge uniformly on $\mathcal{W}^d$ to b and σ, respectively, uniformly in $t \in [0, T]$. Indeed for any $\epsilon > 0$, sufficiently small, all $\phi \in \mathcal{W}^d$,

$$|\sigma_\epsilon(t,\phi)| \leq \|\sigma_\epsilon - \sigma\|_\infty + |\sigma(t,\phi)| \leq (M+\delta)(1 + \sup_{0\leq s\leq t} |\phi_s|)$$

with $\delta > 0$. With the next lemma we shall prove tightness for solutions $\{X^{\epsilon,v}\}_{\epsilon<1}$ of equation (3.2). To exploit the property we use the Kolmogorov Chentsov criterion's (see theorem A.4).

Lemma 3.5. *Let* $(v^\epsilon)_{\epsilon>0} \subset \mathcal{M}^2[0,T]$ *be such that*

$$\int_0^T \|v_s^\epsilon\|^2 \, ds < N$$

for all $\epsilon \in (0,1]$ *and assume A3. Then the family of solutions* $\{X^{\epsilon,v^\epsilon}\}_{\epsilon>0}$ *of equation (3.2) is tight.*

Proof. We need to estimate

$$\mathbb{E}\left[\left|X_t^{\epsilon,v^\epsilon} - X_s^{\epsilon,v^\epsilon}\right|^p\right].$$

Afterwards we have to take p big enough to ensure the conditions of Kolmogorov-Chentsov's criterion. Assume without loss of generality $s < t$. With remark 3.6 in mind we compute

$$\mathbb{E}\left[\int_s^t \overline{M}^p \Big(1 + \sup_{0\le r\le u} |X_r^{\epsilon,v^\epsilon}|\Big)^p \, du\right] \\ \le 2^{p-1}(t-s)\overline{M}^p[T + C(T,N,\overline{M})(1+|x|^p)].$$

Set $H \doteq 2^{p-1}\overline{M}^p[T + C(T,N,\overline{M})(1+|x|^p)]$. Observe that H does not depend on ϵ. The usual computations exploiting the sub-linearity at infinity yield

$$\begin{aligned}
\mathbb{E}\left[|X_t^{\epsilon,v^\epsilon} - X_s^{\epsilon,v^\epsilon}|^p\right] &\le 3^{p-1}(t-s)^{p-1}\mathbb{E}\left[\int_s^t |b_\epsilon(u, X_u^{\epsilon,v^\epsilon})|^p \, du\right] \\
&+ 3^{p-1}\mathbb{E}\left[\left(\int_s^t |\sigma_\epsilon(u, X_u^{\epsilon,v^\epsilon})||v_u^\epsilon| \, du\right)^p\right] + 3^{p-1}\epsilon^{\frac{p}{2}}\mathbb{E}\left[\left|\int_s^t \sigma_\epsilon(u, X_u^{\epsilon,v^\epsilon}) \, dW_u\right|^p\right] \\
&\le 3^{p-1}(t-s)^p H + 3^{p-1}N^{\frac{p}{2}}(t-s)^{\frac{p}{2}}H + 3^{p-1}c_p\epsilon^{\frac{p}{2}}(t-s)^{\frac{p}{2}}H \\
&\le (t-s)^{\frac{p}{2}} \cdot 3^{p-1}H(T^{\frac{p}{2}} + N^{\frac{p}{2}} + \epsilon^{\frac{p}{2}})
\end{aligned}$$

We have proved that the hypotheses of the Kolmogorov-Chentsov's criterion hold if we take coefficients $\alpha = p$ and $\beta = p/2 > 1$ with $p > 2$; hence the family $(X^{\epsilon,v^\epsilon})_{\epsilon>0}$ is tight. □

Hypothesis H2 modified

In the proof of theorem 3.1, hypothesis H2 is only needed to prove that

$$\lim_{n\to\infty} \mathbb{E}[\Psi_t(X^{n,v^n}, v^n)] = 0$$

where, for any given $t \in [0,T]$, $\Psi_t : \mathcal{W}^d \times S_N \to \mathbb{R}$ is defined by

$$\Psi_t(\phi, v) = \left|\phi(t) - x - \int_0^t b(s,\phi_\cdot) \, ds - \int_0^t \sigma(s,\phi_\cdot)v_s \, ds\right| \wedge 1,$$

We show that if we assume A3 and A4 then the same conclusion holds. To this end we use a localization argument.

Given b_ϵ, set

$$b^R_\epsilon(s,\phi_\cdot) \doteq \begin{cases} b_\epsilon(s,\phi_\cdot) & \text{if } \sup_{0\le u\le s}|\phi_u| \le R, \\ b_\epsilon\big(s, \frac{\phi_\cdot}{\|\phi\|_\infty}R\big) & \text{otherwise.} \end{cases}$$

Similarly define $\sigma^R_\epsilon, \sigma^R$ and b^R. It is clear that the functions just defined are globally Lipschitz and bounded. Besides, by assumption A4, $b^R_\epsilon \to b^R$ and $\sigma^R_\epsilon \to \sigma^R$ uniformly on the whole $\mathcal{W}^d$ uniformly in $t \in [0,T]$. Next set

$$\Psi^R_t(\phi, v) = \left|\phi(t) - x - \int_0^t b^R(s,\phi_\cdot)\,ds - \int_0^t \sigma^R(s,\phi_\cdot)v_s\,ds\right| \wedge 1.$$

Observe that if $\sup_{0\le u\le t}|\phi| \le R$ than $\Psi^R_t(\phi,v) = \Psi_t(\phi,v)$. Now consider the family $\{X^{R,n}\}$ of solutions to the equation

$$dX^{R,n}_t = b^R_{\epsilon_n}(t, X^{R,n}_\cdot)\,dt + \sigma^R_{\epsilon_n}(t,X^{R,n}_\cdot)v^n_t\,dt + \sqrt{\epsilon_n}\sigma^R_{\epsilon_n}(t,X^{R,n}_\cdot)\,dW_t,$$

with $X^{R,n}_0 = x$. The same argument as in theorem 3.1 yields

$$\lim_{n\to\infty}\mathbb{E}[\Psi^R_t(X^{R,n}, v^n)] = 0.$$

For any $R > 0$ and for all $n \in \mathbb{N}$ we denote by τ^n_R the exit time of X^{n,v^n} from the ball of center zero and radius R. Then theorem A.2 implies that

$$\mathbb{P}\big(X^{R,n}_t = X^{n,v^n}_t, \text{ for all } t \le \tau^n_R\big) = 1$$

Notice that if $t < \tau^n_R$, then $\Psi_t(X^{n,v^n}, v^n) = \Psi^R_t(X^{R,n}, v^n)$. In fact, if $t < \tau^n_R$, then $X^{R,n}_s = X^{n,v^n}_s$ for all $s \in [0,t]$. Hence, by remark 3.4, we obtain $b(s, X^{R,n}_\cdot) = b(s, X^{n,v^n}_\cdot)$ for all $s \in [0,t]$, and accordingly $\Psi_t(X^{n,v^n}_\cdot, v) = \Psi_t(X^{R,n}_\cdot, v) = \Psi^R_t(X^{R,n}, v^n)$.

Therefore, recalling that $0 \le \Psi_t, \Psi^R_t \le 1$, we have

$$\begin{aligned}\mathbb{E}[\Psi_t(X^{n,v^n}, v^n)] &= \mathbb{E}\Big[1_{t<\tau^n_R}\Psi_t(X^{n,v^n}, v^n)\Big] + \mathbb{E}\Big[1_{t\ge\tau^n_R}\Psi_t(X^{n,v^n}, v^n)\Big] \\ &\le \mathbb{E}\Big[\Psi^R_t(X^{R,n}, v^n)\Big] + \mathbb{P}(t \ge \tau^n_R) \end{aligned} \tag{3.19}$$

We observe that if $\overline{M}$ is the common constant for the growth at infinity, then since $\int_0^T |v^n_s|^2\,ds \le N$ a.s. and by (3.16)

$$\mathbb{P}(t \ge \tau^n_R) = \mathbb{P}\bigg(\sup_{0\le s\le t}|X^{R,n}_t| \ge R\bigg) \le \frac{C(t,N,\overline{M})(1+|x|^2)}{R^2} \le \frac{C}{R^2}$$

Finally, taking the upper limit as $n \to \infty$ on both sides of (3.19), we obtain

$$\limsup_{n\to\infty} \mathbb{E}[\Psi_t(X^{n,v^n}, v^n)] \leq \limsup_{n\to\infty} \mathbb{P}(t \geq \tau_R^n) \leq \frac{C}{R^2}.$$

Since R has been chosen arbitrarily, it follows that

$$\lim_{n\to\infty} \mathbb{E}[\Psi_t(X^{n,v^n}, v^n)] = 0.$$

Accordingly, the job of assumption H2 is carried out by A3 and A4.

3.3 Applications

In this section we show that Schilder's Theorem and the Freidlin-Wentzell estimates are actually special cases of what has just been proved.

Particularly remarkable is the case of small noise diffusions. Observe that if $\bar{b}(t,x)$ and $\bar{\sigma}(t,x)$ are measurable maps from $[0,T] \times \mathbb{R}^d$ into $\mathbb{R}^d$ and $\mathbb{R}^{d\times m}$ respectively, then

$$b(t,\phi_\cdot) \doteq \bar{b}(t,\phi_t) \qquad \sigma(t,\phi_\cdot) \doteq \bar{\sigma}(t,\phi_t)$$

are predictable mappings.

Example 3.1 (Freidlin-Wentzell estimates)**.** Assume that $\bar{b}(t,x)$ and $\bar{\sigma}(t,x)$ are locally lipschitz continuous in x uniformly in t, so that for all $R > 0$ there exists $L_R > 0$ such that

$$|\bar{b}(t,x) - \bar{b}(t,y)| \leq L_R|x-y| \qquad |\bar{\sigma}(t,x) - \bar{\sigma}(t,y)| \leq L_R|x-y|,$$

for all $t \in [0,T]$ and for all $|x|,|y| \leq R$. Furthermore, assume that there exists a constant $M > 0$ such that

$$|\bar{b}(t,x)| \leq M(1+|x|) \qquad |\bar{\sigma}(t,x)| \leq M(1+|x|)$$

for all $t \in [0,T]$ and for all $x \in \mathbb{R}^d$.

It is plain that under these assumptions b and σ defined above satisfy A1-A4. It follows that the processes $\{X^\epsilon\}$ which satisfy

$$dX_t^\epsilon = \bar{b}(t,X_t)\,dt + \sqrt{\epsilon}\bar{\sigma}(t,X_t)\,dW_t$$

with initial condition $X_0^\epsilon = x$ satisfy the Large deviation principle on $\mathcal{W}^d$, with rate function given by

$$I(\phi) = \inf_{\{v\in L^2([0,T];\mathbb{R}^n):\phi_t = x+\int_0^t \bar{b}(s,\phi_s)\,ds + \int_0^t \bar{\sigma}(s,\phi_s)v_s\,ds\}} \frac{1}{2}\int_0^T \|v_t\|^2\,dt$$

whenever $\{v \in L^2([0,T];\mathbb{R}^n) : \phi_t = x + \int_0^t \bar{b}(s,\phi_s)\,dt + \int_0^t \bar{\sigma}(s,\phi_s)v_s\,ds\} \neq \emptyset$, and $I_x(\phi) = \infty$ otherwise.

Remark 3.8. *For $\bar\sigma(\cdot)$ square matrix, and nonsingular diffusion, namely $a(\cdot) \doteq \bar\sigma(\cdot)\bar\sigma(\cdot)^T$ is uniformly definite positive, the preceding formula for the rate function simplifies considerably to*

$$I(\phi) = \frac{1}{2}\int_0^T (\dot\phi_s - \bar b(\phi_s))^T a^{-1}(\phi_s)(\dot\phi_s - \bar b(\phi_s))\,ds$$

if ϕ is absolutely continuous in $[0,T]$ such that $\phi(0) = x$, $+\infty$ otherwise.

Example 3.2 (Schilder's Theorem)**.** If we take $b_\epsilon \equiv b \equiv 0$ and $\sigma_\epsilon \equiv \sigma \equiv 1$, then hypotheses A1-A4 are obviously satisfied. Therefore, the processes $X^\epsilon = \{X_t^\epsilon : t \in [0,T]\}$ which satisfies the stochastic differential equation

$$dX_t^\epsilon = \sqrt{\epsilon}\,dW_t$$

with initial condition $X_0^\epsilon = 0$, satisfy the Large deviation principle on $\mathcal{W}^d$, with rate function given by

$$I(\phi) = \inf_{\{v\in L^2([0,T];\mathbb{R}^n):\phi_t = \int_0^t v_s\,ds\}} \frac{1}{2}\int_0^T \|v_t\|^2\,dt$$

whenever $\{v \in L^2([0,T];\mathbb{R}^n) : \phi_t = \int_0^t v_s\,ds\} \neq \emptyset$, and $I_x(\phi) = \infty$ otherwise.

Observe that the set $\{v \in L^2([0,T];\mathbb{R}^n) : \phi_t = \int_0^t v_s\,ds\} \neq \emptyset$ if and only if ϕ is absolutely continuous in $[0,T]$ and $\phi(0) = 0$; moreover in that case the set has only one point which is $\dot\phi$. We have proved that the rate function is

$$I(\phi) = \frac{1}{2}\int_0^T \|\dot\phi_t\|^2\,dt$$

whenever ϕ is absolutely continuous in $[0,T]$ such that $\phi(0) = 0$, $+\infty$ otherwise.

An other application of theorem 3.2 can be given for systems with memory [see Mohammed and Zhang, 2006].

Let $\bar b : [0,T]\times\mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}^d$ and $\bar\sigma : [0,T]\times\mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}^{d\times m}$ be Borel measurable functions. We make the following assumptions.

(i) The functions $\bar b$ and $\bar\sigma$ satisfy a Lipschitz condition. That is, there exists a constant $L > 0$ such that for all $x_1, x_2, y_1, y_2 \in \mathbb{R}^d$ and all $t \in [0,T]$,

$$|\bar b(s,x_1,y_1) - \bar b(s,x_2,y_2)| \le L(|x_1 - x_2| + |y_1 - y_2|)$$
$$|\bar\sigma(s,x_1,y_1) - \bar\sigma(s,x_2,y_2)| \le L(|x_1 - x_2| + |y_1 - y_2|)$$

(ii) The functions $\bar{b}(\cdot, x, y)$ and $\bar{\sigma}(\cdot, x, y)$ are continuous on $[0, T]$, uniformly in $x, y \in \mathbb{R}^d$, that is,

$$\lim_{s \to t} \sup_{x,y \in \mathbb{R}^d} |\bar{b}(s, x, y) - \bar{b}(t, x, y)| = 0$$
$$\lim_{s \to t} \sup_{x,y \in \mathbb{R}^d} |\bar{\sigma}(s, x, y) - \bar{\sigma}(t, x, y)| = 0$$

Let $0 < \tau < T$ be a fixed delay, and ψ be a given continuous function on $[-\tau, 0]$. Consider, for $t \in [0, T]$, the following stochastic differential delay equation

$$dX_t^\epsilon = \bar{b}(t, X_t^\epsilon, X_{t-\tau}^\epsilon)\, dt + \sqrt{\epsilon}\, \bar{\sigma}(t, X_t^\epsilon, X_{t-\tau}^\epsilon)\, dW_t, \tag{3.20}$$

with the condition

$$X_s^\epsilon = \psi(s) \qquad s \in [-\tau, 0].$$

Denote by $\mathcal{C}_\psi[-\tau, T]$ the set of all continuous functions $\phi : [-\tau, T] \to \mathbb{R}^d$ such that $\phi(t) = \psi(t)$ for all $t \in [-\tau, 0]$. Consider the map $F : L^2[0, T] \to \mathcal{C}_\psi[-\tau, T]$ which takes $v \in L^2[0, T]$ to the unique solution of the equation

$$\begin{aligned} \phi_t &= \phi_0 + \int_0^t \bar{b}(s, \phi_s, \phi_{s-\tau})\, ds + \int_0^t \bar{\sigma}(s, \phi_s, \phi_{s-\tau}) v_s\, ds && t \in [0, T] \\ \phi_t &= \psi_t && t \in [-\tau, 0] \end{aligned}$$

Theorem 3.4. *Assume (i) and (ii). Then the processes $\{X^\epsilon\}$ which satisfy equation* (3.20) *satisfy a Large deviation principle on $C_\psi[-\tau, T]$, with good rate function given by*

$$I(\phi) = \inf_{\{v \in L^2([0,T]; \mathbb{R}^n) : \phi = F(v)\}} \frac{1}{2} \int_0^T \|v_t\|^2\, dt$$

whenever $\{v \in L^2([0, T]; \mathbb{R}^n) : \phi = F(v)\} \neq \emptyset$, and $I(\phi) = \infty$ otherwise.

Proof. We prove the theorem by means of the contraction principle [Dembo and Zeitouni, 1998, Theorem 4.2.1, page 126], accordingly we prove a large deviation principle on $\mathcal{W}^d$ through theorem 3.1 and then we derive the large deviation principle on $C_\psi[-\tau, T]$.

Consider the function Φ from $\mathcal{W}^d$ into $C_\psi[-\tau, T]$ defined by

$$\Phi[\phi](s) \doteq \begin{cases} \psi(s) 1_{[-\tau,0]}(s) + \phi(s) 1_{]0,T]}(s) & \text{if } \phi \in C_{\psi(0)}[0, T] \\ \psi(s) 1_{[-\tau,0]}(s) + \psi(0) 1_{]0,T]}(s) & \text{otherwise.} \end{cases}$$

Set

$$b(t, \phi_\cdot) \doteq \begin{cases} \bar{b}(s, \phi_s, \psi_{s-\tau}) & \text{if } s \in [0, \tau], \\ \bar{b}(s, \phi_s, \phi_{s-\tau}) & \text{if } s \in [\tau, T] \end{cases}$$

In the same manner, define σ. Consider the stochastic differential equation

$$dY_t^\epsilon = b(t, Y_\cdot^\epsilon)\, dt + \sqrt{\epsilon}\,\sigma(t, Y_\cdot^\epsilon)\, dW_t, \tag{3.21}$$

with initial condition $Y_0^\epsilon = \psi(0)$. Observe that if Y^ϵ is the solution of (3.21) and X^ϵ is the solution of (3.20), then $X^\epsilon = \Phi[Y^\epsilon]$ almost surely. This is why we are going to derive a large deviation principle for $\{X^\epsilon\}$ through theorem 3.1.

We show that the functions b and σ enjoy assumptions (A1-A4). Since the coefficients does not depend on ϵ, it suffices to verify A1 and A2. We check the assumptions for b, since the work for σ is analogue.

We start by proving A1. Observe that from (i) we get

$$|\bar{b}(t, x, y)| \le L(|x| + |y|) + |b(t, 0, 0)|.$$

By (ii) it follows that $\sup_{t\in[0,T]} |b(t, 0, 0)| < +\infty$. Now fix $\phi \in \mathcal{W}^d$, if $s \in [0, \tau]$, then

$$\begin{aligned} |b(s, \phi_\cdot)| &= |\bar{b}(s, \phi_s, \psi_{s-\tau})| \\ &\le L(|\phi_s| + |\psi_{s-\tau}|) + |b(s, 0, 0)|. \end{aligned}$$

If $s \in [\tau, T]$, then

$$\begin{aligned} |b(s, \phi_\cdot)| &= |\bar{b}(s, \phi_s, \phi_{s-\tau})| \\ &\le L(|\phi_s| + |\phi_{s-\tau}|) + |b(s, 0, 0)|. \end{aligned}$$

Accordingly, $|b(s, \phi_\cdot)| \le M(1 + \sup_{0\le u\le s} |\phi_u|)$, where M is big enough.

Next we prove A2. Let $\phi, \theta \in \mathcal{W}^d$. If $s \in [0, \tau]$, then

$$\begin{aligned} |b(s, \phi_\cdot) - b(s, \theta_\cdot)| &= |\bar{b}(s, \phi_s, \psi_{s-\tau}) - \bar{b}(s, \theta_s, \psi_{s-\tau})| \\ &\le L|\phi_s - \theta_s| \le L \sup_{0\le u\le s} |\phi_u - \theta_u|. \end{aligned}$$

If $s \in [\tau, T]$, then

$$\begin{aligned} |b(s, \phi_\cdot) - b(s, \theta_\cdot)| &= |\bar{b}(s, \phi_s, \phi_{s-\tau}) - \bar{b}(s, \theta_s, \theta_{s-\tau})| \\ &\le L\Big(|\phi_s - \theta_s| + |\phi_{s-\tau} - \theta_{s-\tau}|\Big) \le 2L \sup_{0\le u\le s} |\phi_u - \theta_u|. \end{aligned}$$

Therefore b is globally Lipschitz continuous, uniformly in $t \in [0, T]$ and with constant $2L$. Since both A1 and A2 are satisfied, the family $\{Y^\epsilon\}$ of solution of equation (3.21) satisfy a large deviation principle on $\mathcal{W}^d$ with good rate function

$$I(\phi) = \inf_{\{v\in L^2([0,T];\mathbb{R}^n):\phi_t=\psi(0)+\int_0^t b(s,\phi_s)\, ds+\int_0^t \sigma(s,\phi_s)v_s\, ds\}} \frac{1}{2}\int_0^T \|v_t\|^2\, dt.$$

Observe that the effective domain of I, $\mathcal{D}_I = \{\phi \in \mathcal{W}^d : I(\phi) < \infty\}$, is contained in $C_{\psi(0)}[0,T]$. Since the map Φ is continuous from $C_{\psi(0)}[0,T]$ into $C_\psi[-\tau,T]$ (and in fact Φ is a continuous bijection), the contraction principle implies that the family $\{X^\epsilon\}_{\epsilon>0}$ satisfies a large deviation principle on $C_\psi[-\tau,T]$ with good rate function

$$\begin{aligned} J(\phi) &= \inf\left\{ I(\theta) : \phi = \Phi[\theta] \right\} = I(\phi_{|[0,T]}) \\ &= \inf_{\{v \in L^2([0,T];\mathbb{R}^n) : \phi_{|[0,T]} = \Gamma_{\psi(0)}(v)\}} \frac{1}{2} \int_0^T \|v_t\|^2 \, dt \\ &= \inf_{\{v \in L^2([0,T];\mathbb{R}^n) : \phi = F(v)\}} \frac{1}{2} \int_0^T \|v_t\|^2 \, dt. \end{aligned}$$

Indeed, $\phi = F(v)$ if and only if $\phi_{|[0,T]} = \Gamma_{\psi(0)}(v)$. □

3.4 Positive and Hölder continuous diffusions

We are now going to establish a large deviation principle in the particular case where the diffusion coefficient σ is positive and Hölder continuous with exponent $\gamma \geq \frac{1}{2}$ and in dimension one. A very well known example in finance in this setting is the so called CEV model, in which the interest rate is driven by the following dynamic

$$dr_t = \alpha(b - r_t)\,dt + \rho r_t^\gamma \, dW_t. \quad r_0 > 0$$

When $\gamma = \frac{1}{2}$, we obtain, in mathematical finance, the Cox-Ingersoll-Ross model (or CIR model) which describes the evolution of interest rates.

The aim of this section is to derive, through a weak convergence approach, Wentzell-Freidlin large deviation estimates for diffusions with coefficients that are neither bounded nor Lipschitz continuous. This problem has been studied with a different approach by Baldi and Caramellino [2011]. Some of the ideas of their proof are contained in this section.

As usual we consider X^ϵ to be the solution to the equation

$$dX_t^\epsilon = b(X_t^\epsilon)\,dt + \sqrt{\epsilon}\sigma(X_t^\epsilon)\,dW_t, \quad X_0^\epsilon = x > 0 \tag{3.22}$$

Moreover we make the following assumption on the coefficients.

R1 The diffusion coefficient $\sigma : \mathbb{R} \to \mathbb{R}^+$ is locally Lipschitz continuous on $\mathbb{R} \setminus \{0\}$, vanishes at 0, and has sub-linear growth at ∞. Moreover we assume that there exists a continuous positive increasing function $\rho(u)$, $u \in [0,\infty)$ such that

$$|\sigma(x) - \sigma(y)| \leq \rho(|x-y|) \quad \forall x, y \in \mathbb{R}$$

and

$$\int_{0+} \rho^{-2}(u)\, du = +\infty$$

R2 The drift $b : \mathbb{R} \to \mathbb{R}$ is locally Lipschitz continuous, has a sublinear growth at ∞ and $b(0) > 0$.

Remark 3.9. *Remark that the above assumptions imply in particular that there exist $\beta > 0$ and a neighborhood of zero such that $b(x) \geq \beta$ and $\sigma(x) \leq \rho(|x|)$.*

The large deviation principle will be proved showing that the hypotheses of theorem 3.1 are satisfied. It is clear that we shall consider

$$\bar{\sigma}(s, \phi_{.}) = \sigma(\phi_s), \quad \bar{b}(s, \phi_{.}) = b(\phi_s)$$

in order to apply Theorem 3.1. The control of the assumptions is not a big deal, since most of the work has already been carried out in the Lipschitz continuous case. Besides, for sake of simplicity we consider only the case $b_\epsilon \equiv b$ and $\sigma_\epsilon \equiv \sigma$ so that H2 holds true. We remark that hypothesis of sub-linear growth at infinity is sufficient to grant H6.

It is clear that H1 is verified: observe that for any fixed $\phi \in \mathcal{W}^1$, $|\bar{\sigma}(s, \phi_{.})| = |\sigma(\phi_s)| \leq M(1 + \|\phi\|_\infty) < \infty$.

Hypothesis H3 deserves a special discussion. It is not plain that the stochastic differential equation has a pathwise unique solution. In fact, this result is the content of a theorem due to Yamada and Watanabe, which we have collected in the appendix (Theorem A.3). By assumptions R1 and R2 it follows that pathwise uniqueness holds. Since the coefficients are continuous, and have sub-linear growth at infinity, there exists a weak solution to (3.22) defined on $[0, T]$ [see Ikeda and Watanabe, 1989, Theorem 2.3 and Theorem 2.4]. Moreover, we remark that pathwise uniqueness implies that any solution is strong (see Revuz and Yor [1999], Theorem (1.7) page 368); thus, the stochastic differential equation (3.22) has a pathwise unique strong solution. Accordingly, hypothesis H3 holds.

Remark 3.10. *Concerning H4, existence (not uniqueness) to the equation* (3.3) *is granted by sub-linear growth at infinity and by the hypothesis of uniform continuity on bounded subsets of $\mathbb{R}$, which comes by R1 and R2, as we have already seen in the case of Lipschitz coefficients (Lemma 3.1). To discuss uniqueness we shall need the following result.*

Proposition 3.1. *Under Assumptions R1 and R2 the integral equation*

$$\phi_t = x_0 + \int_0^t b(\phi_s)\, ds + \int_0^t \sigma(\phi_s) v_s\, ds$$

for $v \in L^2[0, T]$, admits a unique solution for $t \in [0, T]$. Moreover for every $N > 0$ there exists $\eta > 0$ such that $\phi_t \geq \eta$ for any choice of $\|v\|_{L^2} \leq N$.

Proof. We have already discussed existence in remark 3.10. In order to prove that the solution is unique it suffices to verify the last part of the theorem. Indeed, for any fixed $\|v\|_{L^2} \leq N$, take ϕ and ψ solutions to the above integral equation. Since $\phi_t, \psi_t \geq \eta$, they solves the same equation in which b and σ are locally Lipschitz on the whole $\mathbb{R}$, and we already know that in that case uniqueness holds.

For any absolutely continuous path $\phi : [0, T] \to \mathbb{R}$ with a square integrable derivative, let

$$\Gamma_{t_1}^{t_2}(\phi) = \int_{t_1}^{t_2} \mathcal{L}(\phi_t, \dot{\phi}_t)\, dt, \quad \mathcal{L}(\phi_t, \dot{\phi}_t) = \frac{(\dot{\phi}_t - b(\phi_t))^2}{2\sigma(\phi_t)^2}$$

where b and σ are for now supposed to be strictly positive continuous functions on $]0, +\infty[$, and $0 \leq t_1 < t_2 \leq T$. ($\Gamma_{t_1}^{t_2} = +\infty$ possibly). We would like to estimate $\Gamma_{t_1}^{t_2}$ from below uniformly in ϕ. To this end we use a approach of calculus of variations.

Let

$$H(x, p) = \sup_{v \in \mathbb{R}} (vp - \mathcal{L}(x, v)) = \sup_{v \in \mathbb{R}} \left(vp - \frac{(v - b(x))^2}{2\sigma(x)^2} \right)$$

A straightforward computation yields

$$H(x, p) = b(x)p + \frac{1}{2}\sigma(x)^2 p^2.$$

By construction, it is clear that

$$H(x, p) + \mathcal{L}(x, v) \geq vp \tag{3.23}$$

for every v, p and $x > 0$. Next we consider, for $x, x_0 > 0$ the function

$$V(x) = -2 \int_{x_0}^{x} \frac{b(z)}{\sigma(z)^2}\, dz$$

We can easily check that V solves the problem

$$\begin{cases} H(x, w') = 0 \\ w(x_0) = 0 \end{cases}$$

Let ϕ be an absolutely continuous path such that $\phi_{t_1} = x_0$, $\phi_{t_2} = x$ and $\phi_t \in [x, x_0]$ for all $t \in [t_1, t_2]$. Then by (3.23) we get

$$\begin{aligned} V(x) &= \int_{t_1}^{t_2} \frac{d}{ds} V(\phi_s)\, ds = \int_{t_1}^{t_2} V'(\phi_s)\dot{\phi}_s\, ds \leq \\ &\leq \int_{t_1}^{t_2} \left(H(\phi_s, V'(\phi_s)) + \mathcal{L}(\phi_s, \dot{\phi}_s) \right) ds = \Gamma_{t_1}^{t_2}(\phi) \end{aligned}$$

We have proved that for all ϕ absolutely continuous path such that $\phi_{t_1} = x_0$, $\phi_{t_2} = x$ and $\phi_t \in [x, x_0]$ for all $t \in [t_1, t_2]$

$$V(x) \leq \Gamma_{t_1}^{t_2}(\phi)$$

Finally we have all the elements to prove the result in the statement. Let $\bar{x} > 0$ be such that $b(x) \geq \beta > 0$ for some $\beta > 0, \delta > 0$ and for $x \in [0, \bar{x}]$. We can obviously assume also that $\bar{x} < x_0$. Let ξ, $0 < \xi < \bar{x}$. We are going to prove that if ξ is sufficiently small and $\|v\|_{L^2} \leq N$, then $\phi_t > \xi$ for every $t \in [0, T]$. Indeed, otherwise, there would exist two times $t_1 < t_2 \leq T$ such that $\phi_{t_1} = \bar{x}$, $\phi_{t_2} = \xi$ and $\phi_t \in [\xi, \bar{x}]$ for all $t_1 \leq t \leq t_2$. Then

$$\|v\|_{L^2}^2 = \int_0^T \frac{(\dot{\phi}_t - b(\phi_t))^2}{2\sigma(\phi_t)^2}\, dt \geq \int_{t_1}^{t_2} \frac{(\dot{\phi}_t - b(\phi_t))^2}{2\sigma(\phi_t)^2}\, dt.$$

as $b(x) \geq \beta > 0$ and $\sigma(x) \leq \rho(|x|)$ for $x \leq \bar{x}$, by the above computation

$$\begin{aligned}\int_{t_1}^{t_2} \frac{(\dot{\phi}_t - b(\phi_t))^2}{2\sigma(\phi_t)^2}\, dt &\geq -2\int_{\bar{x}}^{\xi} \frac{b(z)}{\sigma(z)^2}\, dz \\ &\geq 2\int_{\xi}^{\bar{x}} \frac{\beta}{\rho^2(z)}\, dz\end{aligned}$$

Now remember that, by assumption

$$\int_{0+} \rho^{-2}(u)\, du = +\infty;$$

thus, pick $\xi > 0$ such that $\beta \int_{\xi}^{\bar{x}} \rho^{-2}(u)\, du > N^2$. This is in contradiction with the assumption $\|v\|_{L^2} \leq N$. Therefore, if $\|v\|_{L^2} \leq N$, then $\phi_t > \xi$ for every $t \in [0, T]$. In particular, the solution ϕ stays away from 0, so that the equation has a unique solution, as the coefficients are assumed to be Lipschitz continuous on $]0, +\infty[$. □

By the above stated proposition it follows that H4 is satisfied. By the same proposition it is not difficult to observe that H5 is satisfied. The map Γ_x which takes $v \in S_N$ to the unique solution of the integral equation

$$\phi_t = x + \int_0^t b(\phi_s)\, ds + \int_0^t \sigma(\phi_s) v_s\, ds$$

coincides with the map defined by replacing σ with a function which is locally Lipschitz on the whole $\mathbb{R}$ and equals σ out of a sufficiently small neighborhood of zero. Indeed, there exists $\xi > 0$ such that, for all $v \in S_N$, $\Gamma_x(v) \geq \xi$. Therefore, Γ_x is continuous from S_N endowed with the weak topology of L^2, by what we have already proved in the case of Lipschitz continuous coefficients.

By assumptions R1 and R2, it follows immediately that H6 is satisfied. Indeed the only thing we need here is that σ and b have sublinear growth at infinity so that we can argue exactly as in the Lipschitz continuous case.

Finally, since the hypotheses for the LDP are satisfied we can state the following.

Theorem 3.5. *Assume R1 and R2. Then the family* $\{X^\epsilon\}$ *of solutions of the stochastic differential equation* (3.1)*, taking values in the Polish space* $C([0,T];\mathbb{R})$*, satisfies the Laplace principle in* $C([0,T];\mathbb{R})$ *with (good) rate function* I_x*, given by*

$$I_x(\phi) = \frac{1}{2}\int_0^T \frac{(\dot{\phi}_t - b(\phi_t))^2}{\sigma^2(\phi_t)} dt$$

whenever $\phi \in AC_x[0,T]$ *and* $I_x(\phi) = \infty$ *otherwise.*

Proof. We have already checked that R1 and R2 imply (H1-H6). Therefore, Theorem 3.1 implies that the family $\{X^\epsilon\}$ of solutions of the stochastic differential equation (3.1), taking values in the Polish space $C([0,T];\mathbb{R})$, satisfies the Laplace principle in $C([0,T];\mathbb{R})$ with rate function I_x, given by

$$I_x(\phi) = \inf_{\{v \in L^2([0,T];\mathbb{R}^n) : \phi_t = \Gamma_x(v)\}} \frac{1}{2}\int_0^T \|v_t\|^2 \, dt$$

whenever $\{v \in L^2([0,T];\mathbb{R}^n) : \phi_t = \Gamma_x(v)\} \neq \emptyset$, and $I_x(\phi) = \infty$ otherwise. Observe that if $\{v \in L^2([0,T];\mathbb{R}^n) : \phi_t = \Gamma_x(v)\} \neq \emptyset$ then, for any u in that set

$$\phi_t = x + \int_0^t b(\phi_s)\, ds + \int_0^t \sigma(\phi_s) u_s \, ds.$$

Thus $\phi \in AC_x[0,T]$. Moreover

$$\dot{\phi}_t = b(\phi_t) + \sigma(\phi_t) u_t.$$

We have already proved that $\phi_t > 0$ for $t \in [0,T]$, thus $\sigma(\phi_t) \neq 0$ and accordingly

$$\frac{\dot{\phi}_t - b(\phi_t)}{\sigma(\phi_t)} = u_t.$$

The last display implies that for all $u \in \{v \in L^2([0,T];\mathbb{R}^n) : \phi_t = \Gamma_x(v)\}$

$$\frac{1}{2}\int_0^T \|u_t\|^2 \, dt = \frac{1}{2}\int_0^T \frac{(\dot{\phi}_t - b(\phi_t))^2}{\sigma^2(\phi_t)} dt$$

Hence the proof is over. □

A
Appendix

Stochastic differential Equations

In this section we collect some definitions and theorems about the stochastic differential equations. In what follows we deal with the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, where $\{\mathcal{F}_t\}$ is supposed to be right continuous and complete.

Definition A.1 (Predictable path functionals)**.** *Let $\mathcal{P}$ be the predictable σ-algebra, that is the σ-algebra generated by the space of the $\mathcal{F}_t$-adapted processes which are left-continuous on $]0, T]$. A process X with values in a measurable space $(E, \mathcal{E})$ is $(\mathcal{F}_t)$-predictable if the map $X : \Omega \times [0, T] \to E$ is measurable with respect to $\mathcal{P}$.*

A function $f : [0, T] \times \mathcal{W}^n \to \mathbb{R}$ is said to be predictable if it is predictable as a process defined on $\mathcal{W}^n$ with respect to $\mathcal{G}_t$. Now let X be a continuous adapted $\mathbb{R}^n$-valued process defined on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, then X takes values in $\mathcal{W}^n$, and if f is predictable, we will write $f(s, X_\cdot)$ or $f(s, X_\cdot(\omega))$ for the value taken by f on the path $s \to X_s(\omega)$ at the time s. It is not difficult to check that $f(s, X_\cdot)$ is $\mathcal{F}_t$-predictable.

It has to be clear that $f(s, X_\cdot(\omega))$ may depend on the whole path up to time s. A very special case is when $f(s, \phi_\cdot) = \sigma(s, \phi(s))$, where σ is defined on $[0, T] \times \mathbb{R}^d$. In that situation $f(s, X_\cdot) = \sigma(s, X_s)$.

Let $b(\cdot, \cdot)$ and $\sigma(\cdot, \cdot)$ be predictable functions from $[0, T] \times \mathcal{W}^d$ to $\mathbb{R}^d$

and to $\mathbb{R}^{d\times m}$ respectively. Fix $x \in \mathbb{R}^d$, and consider the stochastic functional differential equation

$$dX_t = b(t, X_\cdot)\, dt + \sigma(t, X_\cdot)\, dW_t, \tag{A.1}$$

for $t \in [0, T]$, and with initial condition $X_0 = x$.

Definition A.2. *A* strong solution *of the stochastic differential equation* (A.1), *on the given probability space* $(\Omega, \mathcal{F}, \mathbb{P})$, *with respect to the fixed Brownian motion W and initial condition* $x \in \mathbb{R}^d$, *is a process* $X = \{X_t; 0 \le t \le T\}$ *with continuous sample paths and with the following properties:*

(i) *X is adapted to the augmented filtration generated by W;*

(ii)

$$\mathbb{P}\left(\int_0^T |b(s, X_\cdot)| + \|\sigma(s, X_\cdot)\|^2\, ds < +\infty\right) = 1$$

(iii)

$$X_t = x + \int_0^t b(s, X_\cdot)\, ds + \int_0^t \sigma(s, X_\cdot)\, dW_s$$

for all $t \in [0, T]$, $\mathbb{P}$*-almost surely.*

Definition A.3. *Let the drift vector b and the dispersion matrix σ be given. Suppose that, whenever W is a d-dimensional Brownian motion on some* $(\Omega, \mathcal{F}, \mathbb{P})$, *and X, Y are two strong solutions to* (A.1) *relative to W with initial condition* $x \in \mathbb{R}^d$, *then* $\mathbb{P}(X_t = Y_t,\ t \in [0, T]) = 1$. *Under these conditions we say that* strong uniqueness *holds for the pair* (b, σ).

Definition A.4. *A* weak solution *of the stochastic differential equation* (A.1) *is a triple* (X, W), $(\Omega, \mathcal{F}, \mathbb{P})$, $\{\mathcal{F}_t\}$, *where*

(i) $(\Omega, \mathcal{F}, \mathbb{P})$ *is a probability space, and* $\{\mathcal{F}_t\}$ *is a filtration of sub-σ-fields of* $\mathcal{F}$ *satisfying the usual conditions.*

(ii) *X is a continuous,* $\mathcal{F}_t$*-adapted,* $\mathbb{R}^d$ *valued process, W is an r-dimensional* $\mathcal{F}_t$*-Brownian motion and (ii), (iii) conditions of definition A.2 are satisfied.*

Definition A.5. *Suppose that, whenever* (X, W), $(\Omega, \mathcal{F}, \mathbb{P})$, $\{\mathcal{F}\}$ *and* $(\tilde{X}, W)$, $(\Omega, \mathcal{F}, \mathbb{P})$, $\{\tilde{\mathcal{F}}\}$, *are weak solutions to* (A.1) *with common Brownian motion W (relative to possibly different filtrations) on a common probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and with common initial value, i.e.* $\mathbb{P}(X_0 = \tilde{X}_0) = 1$, *the two processes are indistinguishable:* $\mathbb{P}(X_t = \tilde{X}_t,\ t \in [0, T]) = 1$. *We say then that* pathwise uniqueness *hold for equation* (A.1).

Definition A.6. *We say that* uniqueness in the sense of probability law *holds for equation* (A.1) *if, for any two weak solutions* (X, W), $(\Omega, \mathcal{F}, \mathbb{P})$, $\{\mathcal{F}_t\}$, *and* $(\tilde{X}, \tilde{W})$, $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, $\{\tilde{\mathcal{F}}_t\}$, *with the same initial distribution, i.e.*

$$\mathbb{P}(X_0 \in \Gamma) = \tilde{\mathbb{P}}(\tilde{X}_0 \in \Gamma); \quad \forall \Gamma \in \mathcal{B}(\mathbb{R}^d),$$

the two processes X, $\tilde{X}$ *have the same law.*

Theorem A.1. *Suppose that* b *and* σ *satisfy the following assumptions:*

(i) b *and* σ *satisfy a sublinear growth condition. Specifically, there exists* $M > 0$ *such that for all* $s \in [0, T]$, *all* $\phi \in \mathcal{W}^d$,

$$|b(s, \phi_.)| \leq M(1 + \sup_{0 \leq u \leq s} |\phi_u|) \qquad |\sigma(s, \phi_.)| \leq M(1 + \sup_{0 \leq u \leq s} |\phi_u|).$$

(ii) b *and* σ *are locally Lipschitz continuous. Specifically, for any* $R > 0$ *there exists* $L_R > 0$ *such that*

$$|b(s, \phi_.) - b(s, \psi_.)| \leq L_R \sup_{0 \leq u \leq s} |\phi_u - \psi_u|$$

$$|\sigma(s, \phi_.) - \sigma(s, \psi_.)| \leq L_R \sup_{0 \leq u \leq s} |\phi_u - \psi_u|$$

whenever $\sup_{0 \leq u \leq s} |\phi_u \vee \psi_u| \leq R$ *and* $s \in [0, T]$.

Then pathwise uniqueness and existence of strong solutions hold for equation (A.1).

For a proof see for example Theorem (12.1), page 132 of Rogers and Williams [2000].

Theorem A.2. *Suppose that* b_i *and* σ_i, $i = 1, 2$ *satisfy hypotheses (i) and (ii) above and that* X^i *is a solution to*

$$X_t^i = x + \int_0^t b_i(s, X_.^i)\, ds + \int_0^t \sigma_i(s, X_.^i)\, dW_s$$

on some set-up $(\Omega, \{\mathcal{F}_t\}, \mathbb{P}, W_t)$. *Furthermore, assume that*

$$b_1(s, \phi_.) = b_2(s, \phi_.)$$

$$\sigma_1(s, \phi_.) = \sigma_2(s, \phi_.)$$

on $0 \leq s \leq \tau(\phi_.)$, *where* $\tau : \mathcal{W}^d \to \mathbb{R}$ *is an* $\mathcal{G}_t$*-stopping time. Then*

$$\mathbb{P}(X_t^1 = X_t^2, \textit{ for all } t < \tau(X^1))$$

For a proof, see Corollary (11.10) of Rogers and Williams [2000].

In the theorem which follow we give a uniqueness result when the stochastic differential equation of interest is one-dimensional.

Theorem A.3 (Yamada-Watanabe). *Assume that:*

(i) *there exists a positive increasing function $\rho(u)$, $u \in [0, \infty)$ such that*

$$|\sigma(x) - \sigma(y)| \leq \rho(|x - y|) \quad \forall x, y \in \mathbb{R}$$

and

$$\int_{0+} \rho^{-2}(u)\, du = +\infty;$$

(ii) *the drift b is locally Lipschitz continuous.*

Then pathwise uniqueness holds for equation (A.1).

For a proof we refer to Revuz and Yor [1999], Theorem 3.5 page 390.

Tightness

Here we provide a criterion to decide whether a system of d-dimensional continuous processes $\{X_t^\epsilon\}_{t\in[0,T]}$ is tight.

Theorem A.4 (Kolmogorov Chentsov). *Let $\{X_t^\epsilon\}_{t\in[0,T]}$ be a system of d-dimensional continuous processes satisfying the following two conditions.*

1. *the family of initial distributions $\{\mathbb{P}(X_0^\epsilon \in \cdot)\}_{\epsilon>0}$ is tight;*
2. *there exist positive constants α, β and M such that for every $\epsilon > 0$, and for all $t, s \in [0, T]$*

$$\mathbb{E}[|X_t^\epsilon - X_s^\epsilon|^\alpha] \leq M|t - s|^{1+\beta}$$

Then the family $\{X^\epsilon\}_{\epsilon>0}$ is tight as a family of $\mathcal{C}([0, T]; \mathbb{R}^d)$ valued random variables.

For a proof we refer to Klenke [2008, Theorem 21.42 page 473].

Bibliography

P. Baldi and L. Caramellino. General Freidlin-Wentzell Large Deviations and positive diffusions. *Statistics & Probability Letters*, 81(8):1218 – 1229, 2011.

M. Bou and P. Dupuis. A Variational Representation for Certain Functionals of Brownian Motion. *The Annals of Probability*, 26(4):1641–1659, 1997.

A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Stochastic Modelling and Applied Probability. Springer, 2nd edition, 1998. ISBN 9780387984063.

P. Dupuis and R.S. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley series in probability and statistics: Probability and statistics. Wiley, 1997. ISBN 9780471076728.

D. Gilbarg and N.S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Classics in Mathematics. Springer, 2001. ISBN 9783540411604.

N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*. Kodansha scientific books. North-Holland, 1989. ISBN 9784062032315.

I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer, 1991. ISBN 9780387976556.

A. Klenke. *Probability Theory: A Comprehensive Course*. Universitext (1979). Springer, 2008. ISBN 9781848000476.

S. Mohammed and T. Zhang. Large deviations for stochastic systems with memory. *Discrete and continuous dynamical systems-series B*, 6(4):881–893, 2006.

D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Grundlehren der Mathematischen Wissenschaften. Springer, 1999. ISBN 9783540643258.

L.C.G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales*. Number v. 2 in Cambridge Mathematical Library. Cambridge University Press, 2000. ISBN 9780521775939.

S.R.S. Varadhan. Asymptotic probabilities and differential equations. *Communications on Pure and Applied Mathematics*, 19:261–286, 1966.